MODELING CASCADING FAILURES IN COMPLEX NETWORKS

A Dissertation

by

DAN LV

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Tie Liu
Committee Members,	Shuguang Cui
	Le Xie
	Huiyan Sang
Head of Department,	Miroslav M. Begovic

August 2017

Major Subject: Electrical Engineering

ABSTRACT

Large-scale cascading failures can be triggered by very few initial failures, leading to severe damages in complex networks. As modern society becomes more and more networked, there is an increasing requirement of security and reliability of complex networks such as infrastructure networks and cyber networks. In order to design networks which are robust to attacks and enhance the security of the existing networks, this paper studies *load-dependent* cascading failures in random networks consisting of a large but finite number of components. Under a random single-node attack, a framework is developed to quantify the damage at each stage of a cascade. We mainly use probability theory to analyze the cascade process and use simulations to verify our conclusion. In our result, estimations for the fraction of failed nodes are presented to evaluate the time-dependent system damage due to the attack. Furthermore, the analysis reveals a phase transition behavior in the extent of the damage as the load margin grows. That is, the fraction of the damaged components drops from near one to near zero over a slight change in the load margin. The critical value of the load margin and the short interval over which such an abrupt change occurs are derived to characterize the network reaction to small network load variations. Our findings provide design principles for enhancing the network resiliency and provide guidelines for choosing the load margin to avoid a cascade of failures in load-dependent complex networks with practical sizes.

DEDICATION

To my parents, Jianqiang Lyu and Yingping Tai,

and my fiancé, Sichao Jia.

ACKNOWLEDGEMENTS

I want to thank God for giving me this opportunity to study in U.S. and let me graduate with a Ph.D degree in Electrical Engineering from Texas A&M University. This is an amazing journey. When I arrived here with only two baggages and nothing else five years ago, I couldn't image that before long I could harvest such previous treasures: faith, friends, love, family, experiences, degree and career. I am grateful that He guides me to know Him in this Aggieland and shows His love at every step in the way.

I want to thank my advisors, Dr. Shuguang Cui and Dr. Tie Liu, for their help in my work. Dr. Cui offered me this opportunity to study at Texas A&M University. They both taught me how to think critically and write technically. I am always amazed by their brilliance and it is a privilege to work with them.

I am so grateful that God has given me amazing family and friends who always stood by my side and supported me. I would like to especially thank the senior Ph.D students and postdoctors in our group, Di li, Hang Li, Amir Salimi, Armin Banaei, Li Xu and Ali Eslami. They always guided me in my research path selflessly and shared their valuable life experience with me. I also want to thank my officemates Yanyan Zhang, Han Zhang and Qilian Yu. They always supported me and made my life joyful.

Finally, I would like to thank my committee members, Dr. Le Xie and Dr. Huiyan Sang, for their time and support on my preliminary exam, defense and the dissertation. Their valuable comments helped me a lot in improving the quality of my work.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professor Tie Liu (advisor), Shuguang Cui and Le Xie of the Department of Electrical and Computer Engineering and Professor Huiyan Sang of the Department of Statistics.

All work for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by a Research Assistantship from the Department of Electrical and Computer Engineering, Texas A&M University. This work was made possible in part by DoD with grant HDTRA1-13-1-0029.

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE | Page

1. INTRODUCTION*

As modern society becomes more and more networked, there is an increasing requirement of security and reliability in complex networks such as telecommunication systems and the Internet. A wide-spread cascading failure is a serious threat to such systems [1,19]. Starting with a small disturbance, cascading failures may lead to a complete or major network collapse. Despite an extensive effort to study the properties of cascading failures in complex networks, this field is still under-explored, especially for the cases of finite-size networks. Thus, it is crucial to investigate the properties of cascading failures for such cases to improve the operation security and reliability.

In many real-life networks such as transportation networks and communication networks, each node bears some load, which can be redistributed locally to its neighbors. The loads of the nodes can be a particular material or abstract information [12]. In normal conditions, each node maintains a load below its capacity. When some initial failures occur (due to attacks or internal failures), the failed nodes are removed, and their loads will be redistributed to their neighbors. If the neighbors then become overloaded, they will fail too, and their loads will be redistributed further. Such a process may lead to a cascading failure.

Many researchers have studied this area from different aspects and found valuable results. From a graph model perspective [7], Motter and Lai [13], and Crucitti et al. [2, 8] both adopted betweenness [8, 16] to model the load on a component. In [13], Motter and Lai found that heterogeneous topologies are more vulnerable to attacks. In [8] and [2], the authors used Erdös-Rényi random graphs to study the networks where the load of each

*Reprinted from [11] "Load-Dependent Cascading Failures in Finite-Size Erdös-Rényi Random Networks" by D. Lv and A. Eslami and S. Cui, 2017, Network Science and Engineering, IEEE Trans on (TNSE), Volume 4, Page 129-139. Copyright 2017 by IEEE.

node is proportional to its betweenness. They showed that in such models an intended attack on the most heavily loaded node can collapse the entire network. Load-dependent cascades have been studied in [3, 17, 18]. In [3], the authors considered a network of identical nodes and applied a Poisson branching process to study cascading failures. In [17] and [18], Wang and Rong assumed that both initially-assigned load and redistributed load are proportional to the node's degree. They concluded that in such a setting, an attack on the most heavily loaded node might not always be the most destructive one. The majority of the existing analytical work focuses on cascading failures in asymptotically large networks. However, finite-size networks are of paramount importance in both theoretical analysis and real life. For such networks, the analytical approach has been mostly overlooked (mainly due to technical difficulties), limiting the results to only simulations [2, 8].

In this paper, we develop a framework to study cascading failures in a network represented by a finite-size Erdös-Rényi (ER) random graph [4, 6]. In such a network, each node carries a certain amount of load, and maintains a *load margin* up to which it can tolerate some extra load. If a node becomes overloaded, it fails and its load is redistributed to its neighbors. We adopt the ER random graph as the topology model since such model is homogeneous by construction, and it has the advantage of being mathematically tractable given many important properties [15]. While the ER random graph itself is a basic model, many of its variants with non-Poisson distributions are used to model some real-world networks, such as the Internet and collaboration graphs [14, 15]. A thorough understanding of such complicated networks requires a strong knowledge of the fundamental ER model.

Furthermore, we focus on random single-node attacks and study the propagation of failures. We assume that the initial load at each node is proportional to its degree, which is also a fraction of its capacity, leaving some load margin. Inspired by the nature of load shifting in many real networks such as the transportation network [2, 10, 13], the load redistribution upon failure is assumed to be in proportion to the neighbors' capacities. This paper proposes several novel approaches to quantitatively analyze such cascade model. For example, we partition the overall node set into several subsets according to their potential failing time. In the proposed framework, the main contributions of this paper are:

- Step-by-step analysis of damages at each stage of a potential cascade: To quantify the severity of a cascade, we use the fraction of failed nodes, which we denote as *failure ratio*. We provide a method to calculate the failure ratio at each step of the potential cascade. We also estimate the time when the cascade reaches a steady state. Our results provide insights into choosing the right value of the load margin such that a cascade of failures can be avoided.

- Threshold behavior of the collateral damage: Numerical results show that the failure ratio drops from near 1 to near 0 over a very short interval of the load margin. We find the interval within which such phase transition occurs and derive the critical value of the load margin at which the abrupt change of the failure ratio takes place. The phase transition interval along with the critical value of the load margin characterizes the network reaction to a random single-node attack.

The rest of this paper is organized as follows. Section 2 introduces the system model and defines the notations. A step-by-step analysis of the average failure ratio is provided in Section 3. The phase transition in the average failure ratio is explored in Section 4. Section 5 concludes the paper.

2. LOAD-BASED CASCADE MODEL*

In this section, we describe our model of the load-based cascading failure. We also discuss the topology and attack models. Table 2.1 lists all the notations used in this paper.

2.1 Topology Model

Consider an ER random graph $G(n, p)$, where the graph has n nodes and every two nodes are connected with probability p independent of the other pairs [6]. The overall node set is denoted as V. During each step t of a cascade (rigorously defined later), some nodes are failed (dead), and some are still functioning (alive). For any node $v \in V$, we use $N_t(v)$ to denote the set of functioning neighbors of v at the beginning of step t. Particularly, $N_0(v)$ represents the initial neighbor set of v before any failures. The ***degree*** of node v, denoted by $k(v)$, is the number of initial neighbors of v, i.e., $k(v) = |N_0(v)|$, where $|\cdot|$ is the cardinality of a set. Each node v has a *load* $L_t(v)$ at the beginning of step t. We assume that the initial load of each node is proportional to its degree [17, 20]. To simplify the analysis, we set the proportion scaling factor as 1, such that $L_0(v) = k(v)$. Each node v has a time-invariant *capacity* $C(v)$, which is the maximum load that it can handle. Furthermore, $C(v)$ is assumed to be proportional to $L_0(v)$, i.e.,

$$C(v) \ = \ \alpha L_0(v) \ = \ \alpha k(v), \tag{2.1}$$

where $\alpha \geq 1$ is the *tolerance parameter (load margin)* of the network, the same across all the nodes. A node will fail if the load exceeds its capacity.

---

*Reprinted from [11] "Load-Dependent Cascading Failures in Finite-Size Erdös-Rényi Random Networks" by D. Lv and A. Eslami and S. Cui, 2017, Network Science and Engineering, IEEE Trans on (TNSE), Volume 4, Page 129-139. Copyright 2017 by IEEE.

Table 2.1: List of Notations. Reprinted from [11].

Notation	Description
$G(n, p)$	ER random graph with n nodes and connectivity probability p
V	Set of all nodes
$\lvert \cdot \rvert$	Cardinality of a set
$k(\cdot)$	Initial degree
$N_t(\cdot)$	Functioning neighbor set at the beginning of step t
$L_t(\cdot)$	Load at the beginning of step t
α	Tolerance parameter
$C(\cdot)$	Capacity
$\Delta L_t(v, u)$	Redistributed load from node v to node u at step t
a	Initially attacked node
V_d	Set of nodes whose shortest distance from the initially attacked node is d
$\hat{V}_t$	Set of target nodes at step t
F_t	Set of failed nodes during step t
T_t	Set of all failed nodes by the end of step t
f_t	Failure ratio by the end of step t
μ	Conditional mean of $k(V_1)$ given $\lvert V_1 \rvert$
σ^2	Conditional variance of $k(V_1)$ given $\lvert V_1 \rvert$
$\hat{\mu}$	Conditional mean of $L_0(\hat{V}_t)$ given $\lvert \hat{V}_t \rvert$
$\hat{\sigma}^2$	Conditional variance of $L_0(\hat{V}_t)$ given $\lvert \hat{V}_t \rvert$
$\tilde{\mu}$	Conditional mean of $L_t(F_{t-1})$ given $\lvert \hat{V}_t \rvert$
$\tilde{\sigma}^2$	Conditional variance of $L_t(F_{t-1})$ given $\lvert \hat{V}_t \rvert$
$\alpha_{c,t}$	Critical value of α at step t
$[\alpha_{l,t}, \alpha_{h,t}]$	Threshold interval of α at step t

2.2 Attack and Contagion Models

In this paper, we focus on the case of a random single-node attack, where the initially attacked node is a single randomly chosen node in the network. After the initial attack, we divide the overall cascade process into time slots called *steps*. During each step, the latest failed nodes redistribute their loads to their functioning neighbors. If any of the neighbors fail, they will redistribute their loads at the next step. The load redistribution mechanism is described as follows. Assume that node a is attacked initially (during step $t = 0$) and fails. Then at step $t = 1$, the neighbors of a, i.e., nodes in $N_1(a)$, will receive some redistributed load, with node a and its adjacent links being removed from the network. Given a node $u \in N_1(a)$, $\Delta L_1(a, u)$ denotes the redistributed load received by node u from node a at step 1, where $\Delta L_1(a, u)$ is proportional to the capacity of u, and is given as

$$\Delta L_1(a, u) \triangleq L_1(a) \frac{C(u)}{\sum_{i \in N_1(a)} C(i)}. \tag{2.2}$$

The load of u at the beginning of step 2 can be obtained as $L_2(u) = L_1(u) + \Delta L_1(a, u)$. This redistribution rule makes sense since nodes with larger capacities bear larger absolute load margins. After redistributing the load of a, some nodes in $N_1(a)$ may fail and be removed from the network, and redistribute their total loads to their functioning neighbors at step $t = 2$ in a similar fashion. We define the *steady state* as the step when the failure propagation stops in the network. Let F_t, $t \geq 1$, denote the set of nodes fail during the step t. The set of accumulated failed nodes by the end of step t is denoted by T_t, where $T_t = a + \cup_{i=1}^{t} F_i$. We define the *failure ratio* f_t at step t as $|T_t|$ divided by the network size n, i.e., $f_t = |T_t|/n$, such that $f_t \in [0, 1]$. This paper uses the average value of the failure ratio taken over all random realizations of the ER graph and all random single-node attacks to quantify the extent of the damage caused by a single-node attack.

We define the collective degree for a set $W \subseteq V$ as the summation of the degrees of

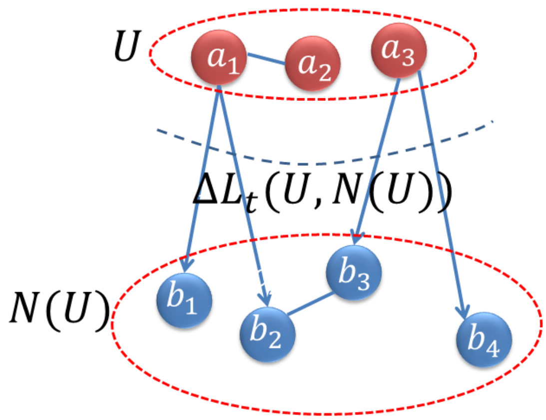

Figure 2.1: An illustration of neighbors of a set U and the load redistribution between two sets. Reprinted from [11].

the nodes in W, given by $k(W) \triangleq \sum_{v \in W} k(v)$. Similarly, the aggregate capacity and load of set W at step t are defined as $C(W) \triangleq \sum_{v \in W} C(v)$ and $L_t(W) \triangleq \sum_{v \in W} L_t(v)$, respectively. The load redistributed at step t from a set W to a disjoint set U is denoted by $\Delta L_t(W, U) \triangleq \sum_{\forall v \in W, \forall u \in U} \Delta L_t(v, u)$. The neighbor set union of set W at step t is given by $N_t(W) \triangleq (\cup_{v \in W} N_t(v)) \setminus W$. An example of the neighboring set and the load redistribution between two sets are illustrated by Fig. 2.1. Utilizing these notations, we can formally present the load redistribution rule at an arbitrary step t as follows. For $t = 1, 2, \cdots$, $\forall v \in F_{t-1}$, and $\forall u \in N_t(v)$, we have

$$\Delta L_t(v, u) \triangleq L_t(v) \frac{C(u)}{C(N_t(v))} = L_t(v) \frac{k(u)}{k(N_t(v))}. \tag{2.3}$$

2.3 Partition of Nodes

To visualize the step-by-step cascade of failures, this paper proposes a partition which can be applied to the overall node set V. According to the partition, set V is divided into a

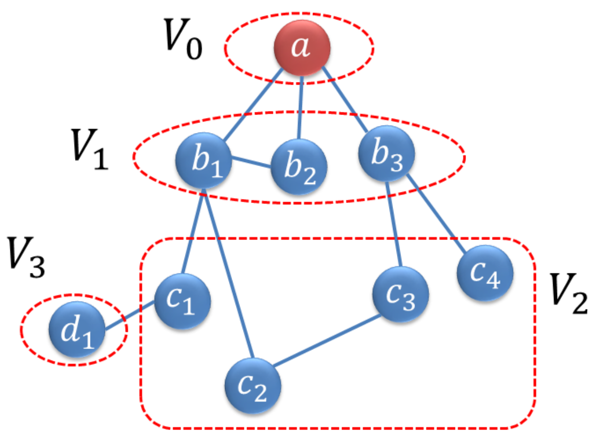

Figure 2.2: Partition example: V is divided into $V_0 = \{a\}, V_1 = \{b_1, b_2, b_3\}, V_2 = \{c_1, c_2, c_3, c_4\}, V_3 = \{d_1\}$. Reprinted from [11].

sequence of distinct subsets $V_0, V_1, V_2, \cdots$, where $V_0 = \{a\}$ is the initially attacked node. For $d \geq 1, V_d$ denotes the set of nodes whose shortest distance (number of hops) to a is d, measured over the initial topology. Note that the minimum number of steps that it takes a cascade to reach to V_d is d. In Fig. 2.2, we demonstrate such a partition, from which the following relationship can be observed between V_d, $d \geq 0$, and their original neighbor sets:

$$V_d = \begin{cases} \{a\} & d = 0 \\ N_0(V_{d-1}) & d = 1 \\ N_0(V_{d-1}) \backslash V_{d-2} & d > 1 \end{cases} . \tag{2.4}$$

These are important properties to be utilized in our analysis. Obtaining $E[|V_d|]$, the average size of V_d, will simplify the analysis in Section 3. The following lemma gives an approximation of $E[|V_d|],\ d \geq 1$, for random single-node attacks in ER random graph.

Lemma 2.3.1 Consider a random single-node attack applied to $G(n,p)$. Let node a be the attacked node, and node e be an arbitrary node in $V\backslash\{a\}$. Let P_d be the probability that *the shortest path from* e *to* a *has length* d*;* $\Pr\{B_d\}$ *be the probability that at least one path from* e *directly through a node in* $V \backslash \{a \cup e\}$ *to* a *has a length which is less than or equal to* d*. Then* $E[|V_d|], d \geq 1$*, the average size of* V_d*, is simply*

$$E[|V_d|] \;=\; (n-1)P_d,$$

where $P_d, d \geq 1$*, can be obtained recursively as*

$$\begin{aligned} P_1 \;&=\; p, \\ P_2 \;&=\; (1-p)(1-(1-p^2)^{n-2}), \\ P_d \;&=\; (1-p)\Pr\{B_d\} - \sum_{j=2}^{d-1} P_j,\; d > 2. \end{aligned}$$

In the numerical calculation, we assume that $\Pr\{B_d\}, d > 2$*, can be approximated recursively as*

$$\Pr\{B_d\} \;\approx\; 1 - ((1-p) + p \cdot (1 - \sum_{j=1}^{d-1} P_j))^{n-2}.$$

The proof of Lemma 2.3.1, including the discussion of the approximation, is provided in the appendix. To see the accuracy of the results in Lemma 2.3.1, some simulations were conducted with different values of n and p. It can be found that Lemma 2.3.1 always yields accurate approximations of $E[|V_d|]$. One example is shown in Fig. 2.3, in which the theoretic estimates were obtained from Lemma 2.3.1, and the simulation results were averaged from $8,000$ random realizations of $G(150, 0.05)$. We can see that the two curves are almost overlapping, such that the approximations given in Lemma 2.3.1 are highly accurate.

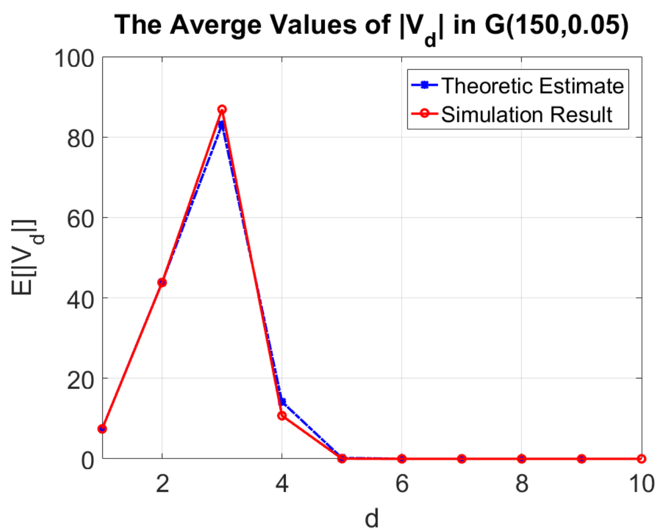

Figure 2.3: Comparison between the theoretic estimates and the simulation results of $E[|V_d|]$ in $G(150, 0.05)$. The theoretic estimates match the simulation results. Reprinted from [11].

3. ESTIMATION OF THE AVERAGE FAILURE RATIO*

Recall that the failure ratio f_t at step t is defined as the total casualties $|T_t|$ divided by the network size n, i.e., $f_t = |T_t|/n$. We use failure ratio to quantify the damage caused by the attack. However, since both the topology (ER random graph) and the location of the initial attack are random, the failure ratio would be a random variable. Thus, we focus on the average value of this random variable taken over all realizations of the random topologies and random single-node attacks. In this section, we first state a few properties that will be utilized in the analysis, and then provide a step-by-step estimate of the average failure ratio. At the end, we verify the accuracy of our estimations via simulations.

3.1 Properties of Load-based Cascade Model

Here are a few preliminary results useful to quantify the average failure ratio. The following lemma and corollary address the first step of load redistribution after the attack. They are particularly useful in early assessment of the attack impacts.

Lemma 3.1.1 Consider an attack on an arbitrary node a in the network $G(n,p)$. Also, consider the partition in Section 2.3 applied to the network. After the load redistribution at step 1, the nodes in V_1 either all fail or all survive.

proof 1 When node a fails, each node $u \in V_1$ receives some additional load given by (2.3). The initial load of u and its capacity are given as $k(u)$ and $\alpha k(u)$, respectively (see (2.1) and the explanation above it). Dividing the total load of u after the redistribution by its

*Reprinted from [11] "Load-Dependent Cascading Failures in Finite-Size Erdös-Rényi Random Networks" by D. Lv and A. Eslami and S. Cui, 2017, Network Science and Engineering, IEEE Trans on (TNSE), Volume 4, Page 129-139. Copyright 2017 by IEEE.

capacity, we have

$$\frac{L_1(u)+\Delta L_1(a,u)}{C(u)} = \frac{k(u)+L_1(a)\frac{k(u)}{k(V_1)}}{\alpha k(u)}, \\ = \frac{k(V_1)+k(a)}{\alpha k(V_1)}, \tag{3.1}$$

which is the same for all nodes in V_1. Therefore, at the first step either all nodes in V_1 fail (i.e., $\frac{L_1(u)+\Delta L_1(a,u)}{C(u)} > 1$) or all of them survive (i.e., $\frac{L_1(u)+\Delta L_1(a,u)}{C(u)} \leq 1$).

Corollary 3.1.1 Consider an attack on an arbitrary node a in the network $G(n,p)$. If the degree constraint

$$k(a)+k(V_1) > \alpha k(V_1) \tag{3.2}$$

holds, all the nodes in V_1 fail at step $t=1$, i.e., $F_1 = V_1$. Otherwise, there are no failures occurring at step $t=1$ or later. In particular, when $\alpha \geq 2$, the degree constraint of (3.2) cannot be satisfied and there would be no cascade of failures.

proof 2 According to the proof of Lemma 3.1.1, for each node $u \in V_1$, the ratio of its load to its capacity by the end of step 1 is given as $\frac{k(V_1)+k(a)}{\alpha k(V_1)}$, which is the same for all the nodes in V_1. Thus, all the nodes in V_1 become overloaded only if this ratio is greater than 1, i.e., $k(a)+k(V_1) > \alpha k(V_1)$. However, the condition $k(V_1) < \frac{k(a)}{\alpha - 1}$ cannot be satisfied when $\alpha \geq 2$ since the aggregate degree of the set V_1, i.e., $k(V_1)$, is always greater than or equal to the degree of the node a, $k(a)$.

Lemma 3.1.1 and Corollary 3.1.1 are direct consequences of our assumptions that the capacity of a node is proportional to its initial load and the load of a failed node is redistributed to its neighbors according to their capacity. There are two important remarks as follows.

Remark 3.1.1 Lemma 3.1.1 and Corollary 3.1.1 hold for each realization of random topology with an arbitrary single-node attack, not only on the average level.

Remark 3.1.2 If $\alpha \leq 1$, there is no tolerance for any extra load from the initially attacked node, and all the nodes will fail as a result of the attack. On the other hand, if $\alpha \geq 2$, Corollary 3.1.1 suggests that the failure would not go beyond the attacked node, resulting in $f_t = 1/n$ for $t \geq 1$. Hence, there is no uncertainty about the network reaction to a random single-node attack when $\alpha \leq 1$ or $\alpha \geq 2$. In the rest of the paper, we will focus on the network reaction with $\alpha \in (1, 2)$.

3.2 Step 1 of Cascade

In the sequel, we present a step-by-step estimate of the average failure ratio that is obtained recursively. The following theorem gives such an estimate after the first step of load redistribution, which will be used to later derive the failure ratio at future steps.

Theorem 3.2.1 Consider a random single-node attack applied to $G(n, p)$. We assume the conditional distribution of $k(V_1)$ given $|V_1| = x$ is approximately normal with mean μ and variance σ^2, where

$$\begin{aligned} \mu &= x + x(x-1)p + (n-x-1)xp, \\ \sigma^2 &= (2x(x-1) + (n-x-1)x)p(1-p). \end{aligned}$$

Then $E[f_1]$, i.e., the average failure ratio at step 1, can be approximated as

$$E[f_1] \approx \frac{1}{n}\left(1 + \sum_{x=1}^{n-1} x \cdot \binom{n-1}{x} p^x (1-p)^{n-1-x} \Phi(\frac{\frac{x}{\alpha-1} - \mu}{\sigma})\right),$$

where $\binom{n-1}{x} = \frac{(n-1)!}{x!(n-x-1)!}$ *is the binomial coefficient, and* $\Phi(\cdot)$ *is the cumulative distribution function (CDF) for the standard normal distribution.*

The proof of Theorem 3.2.1, including the justification of the assumption used in Theorem 3.2.1, is provided in the appendix.

3.3 Future Steps of Cascade

The following theorem provides an estimate for the average failure ratio after the first step, whose proof can be found in the appendix. The assumptions and approximations of Theorem 3.3.1 are justified in the proof.

Theorem 3.3.1 Consider a random single-node attack applied to $G(n, p)$*. We assume,*

1. *We only consider the failures propagating in the forward direction; i.e., at step* t*, only the nodes in* $V \setminus \cup_{d=0}^{t-1} V_d$ *are considered as potential nodes to fail.*

2. *The set* F_{t-1} *is considered as a large virtual node that redistributes its load to its alive neighbors at step* t *with the rule defined in (3).*

3. n *is large enough such that the variance of* $|\hat{V}_t|$ *is small and* $|\hat{V}_t|$ *can be approximated by* $E[|\hat{V}_t|]$*.*

4. $E[|\hat{V}_t| \mid |F_{t-1}| = E[|F_{t-1}|]]$ *is applied to approximate* $E[|\hat{V}_t|]$*.*

5. *Given* $|\hat{V}_t| = E[|\hat{V}_t|]$*,* $L_t(F_{t-1})$ *and* $(\alpha - 1)L_0(\hat{V}_t)$ *are independent and approximately normal.* $L_t(F_{t-1})$ *has conditional mean* $\tilde{\mu} = E[|T_{t-1}|](n-1)p$ *and unknown conditional variance* $\tilde{\sigma}^2$*.* $(\alpha - 1)L_0(\hat{V}_t)$ *has conditional mean* $\hat{\mu} = (\alpha - 1)(n - 1)E[|\hat{V}_t|]p$ *and conditional variance* $\hat{\sigma}^2 = (\alpha - 1)^2(n - 1)E[|\hat{V}_t|]p(1 - p)$*.* $\Phi(\frac{\tilde{\mu}-\hat{\mu}}{\hat{\sigma}})$ *is applied as an approximation of* $\Pr\{L_t(F_{t-1}) > (\alpha - 1)L_0(\hat{V}_t) \mid |\hat{V}_t| = E[|\hat{V}_t|]\}$*.*

Then an estimate of the average failure ratio $E[f_t]$ for step $t \geq 2$ is obtained recursively as

$$E[f_t] \approx \frac{1}{n}\Phi(\frac{\tilde{\mu}-\hat{\mu}}{\hat{\sigma}})E[|\hat{V}_t|] + E[f_{t-1}],$$

where

$$\begin{aligned} E[|\hat{V}_t|] &= (n - \sum_{d=0}^{t-1} E[|V_d|]) \cdot (1-(1-p)^{E[|F_{t-1}|]}), \\ E[|T_{t-1}|] &= nE[f_{t-1}], \\ E[|F_{t-1}|] &= n(E[f_{t-1}] - E[f_{t-2}]), \end{aligned}$$

$E[|V_0|] = 1$ *by definition and* $E[|V_d|],\ d \geq 1$ *are given by Lemma 2.3.1.*

Theorem 3.3.1 provides a recursive method to obtain an estimate of the average failure ratio for all the steps after step 1. Approximations are made to derive the above theorem. One of the approximations, the forward-propagation approximation, will be discussed in the following subsection. Cascade after step 1 is a very complicated process, which involves a large number of random variables. Obtaining an exact closed-form solution is mathematically difficult and computationally complex. In contrast, this theorem takes a computationally manageable expression and provides good accuracy in the same time, which is a simple yet effective approach towards calculating the damage due to a cascade.

3.4 Forward-Propagation Approximation

In the previous analysis, we proposed a step-by-step approximation of the average failure ratio in a cascade event. For mathematical tractability, the approximation only counts the failures propagating in the forward direction, which we denote by "forward propagation". In other words, at each step t of the cascade, we only consider failures caused by V_{t-1} in V_t. This excludes any failures due to the load redistribution within V_{t-1}, and any failures in V_{t-2} caused by V_{t-1} (here denoted as "backward propagation"). Let us explain why the impact of backward propagation is negligible. Recall that Lemma 3.1.1 states that, the nodes in V_1 either all fail or all survive at the first step. Our analysis along with numerical results asserts that this fact *almost* holds for the future steps as well. That is, at each step t, either "almost" all the nodes in V_t die or "almost" all of them survive. Our analysis in Section 4.1 (Theorem 4.1.2) shows that if V_1 fails for a given α, then the chance of $V_2, V_3, \cdots$ failing in the forward failure propagation is quite high, leaving almost no nodes to be failed in the backward propagation. This also results in a *threshold behavior* in the average failure ratio with respect to the load margin. That is, the average failure ratio stays close to either 0 or 1 for almost every value of the load margin while changing from near 1 to near 0 over a very short interval of the load margin. This phase transition is further illustrated by our numerical results in the next subsection and is scrutinized in Section 4.

3.5 Numerical Validation

In the derivation of the step-by-step estimates for $E[f_t]$, several approximations were applied to keep the results tractable. To verify the accuracy of our estimations, a comparison of the estimated average failure ratios given by Theorems 3.2.1 and 3.3.1 against the simulation results is shown in Fig. 3.1. Two topologies, $G(100, 0.05)$ and $G(400, 0.01)$, have been adopted to compute the average failure ratios caused by random single-node attacks. In the simulation results, the empirical average failure ratios were calculated via $8,000$ ER random graph realizations, and 500 experiments with random single-node attacks in each topology, following the load-based failure propagation model defined in Section 2. Fig. 3.1 shows the comparisons at multiple steps of the cascade. The results make it clear that despite several simplifying approximations, our estimates follow the true

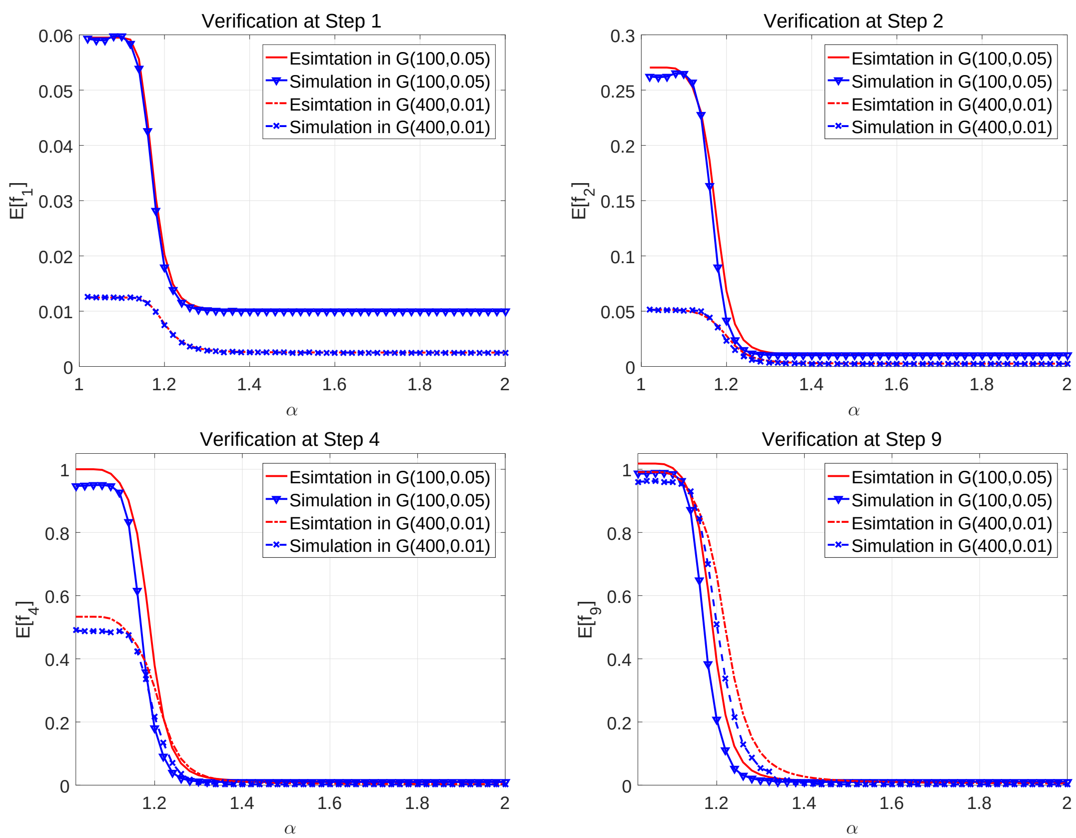

Figure 3.1: Comparison between the estimated average failure ratios and the simulation results in $G(100, 0.05)$ and $G(400, 0.01)$ under random single-node attacks. Estimates for step 1 were given by Theorem 3.2.1, and estimates for steps 2, 4 and 9 were given by Theorem 3.3.1. Reprinted from [11].

(simulated) values very closely. We examined networks of different sizes, and found that the approximation accuracy is acceptable as long as n is reasonably large, e.g., $n \geq 20$. As seen in Fig. 3.1, a random single-node attack may still cause serious damages to the network, particularly when α (load margin) is relatively small. Another important observation is the quick drop in the size of such damages as the load margin grows. Such a *phase transition* phenomenon will be explored in Section 4.

As an illustration of the real-world networks, some simulations were conducted in a Facebook network [9] with $n = 233$ total nodes and $k(V) = 6384$ total degrees. The step-by-step comparison between the average failure ratios in the Facebook network and in ER random graph under random single-node attacks is shown in Fig. 3.2. The ER random graph was generated with the same size n and connecting probability ($p = 6384/n/(n-1) = 0.129$) as the Facebook network. In Fig. 3.2, the simulation results in the Facebook network were obtained from $2,000$ experiments with random single-node attacks in the given network, while the simulation results in ER random graph were obtained from $2,000$ realizations of ER random graph, and 200 experiments with random single-node attacks in each topology.

Furthermore, a comparison between the theoretic estimations of the average failure ratios and the simulation results in the Facebook network, both at steady state, under different α values is shown in Fig. 3.3. The estimated values were calculated by Theorems 3.2.1 and 3.3.1 with setting $n = 233, p = 0.129$. The simulation results were obtained similarly with Fig. 3.2 and the values were recorded at the steady state. Although the Facebook network's power law degree distribution is different from the ER random graph's binomial degree distribution, our model still provides a close trace of the average failure ratio. This example indicates that our model not only makes an important contribution to generic complex network analysis, but also can provide a basis reference for real-world networks. This is because that our model captures the impact of a network's average degree, which

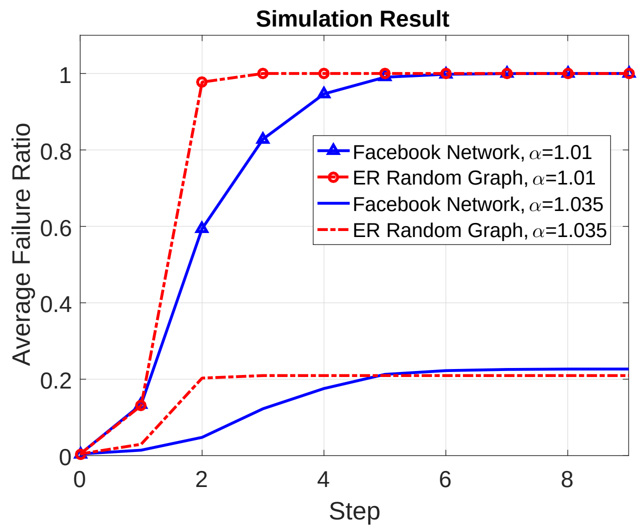

Figure 3.2: Step-by-step comparison between the average failure ratios in the Facebook network and in ER random graph $G(223, 0.129)$ under random single-node attacks. Two α values were tested: $\alpha = 1.01$ and $\alpha = 1.035$. The average failure ratios in both ER random graph and the Facebook network have similar trend and approximately equal values at the steady state. Reprinted from [11].

plays an important role in a cascade of failures.

3.6 Discussion of Other Attack Methods

In the previous subsection, we have shown that the simulation results in ER random graph under random single-node attacks, which validated the analysis. Here we present a brief discussion on other attack methods, including random multiple-node attack and targeted attack on the node with the largest degree, to gain more insight of the proposed model and the analytical results. A comparison between the above two attack methods and random single-node attack is shown in Fig. 3.4. Two different α values were tested: $\alpha = 1.15$ and $\alpha = 1.2$. The simulation results were obtained from $2,000$ realizations of ER graph $G(100, 0.05)$, and 200 experiments with random single-node attacks in each topology.

According to Fig. 3.4, the average failure ratios on the steady state are different under

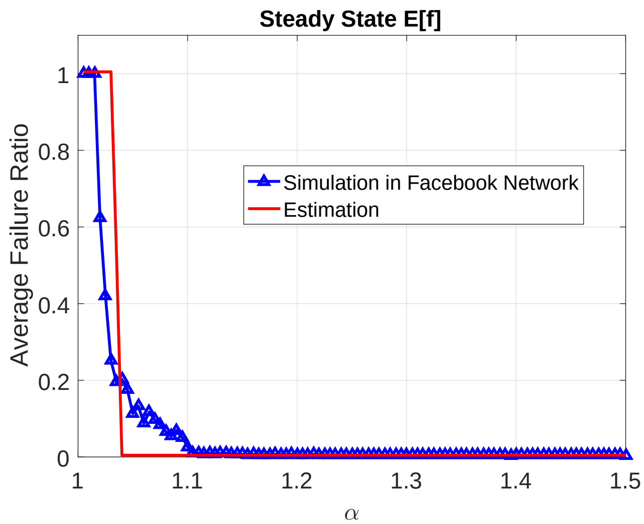

Figure 3.3: Comparison of the estimated average failure ratios against the simulation results in the Facebook network under different α values. The estimation results are given by Theorems 3.2.1 and 3.3.1, which accurately capture the phase transition phenomenon of the average failure ratio over α. Reprinted from [11].

three attack methods. A random five-node attack is always more devastating than a random single-node attack under both two α values. A targeted attack on the node with the largest degree is more destructive than a random single-node attack when $\alpha = 1.15$. In contrast, when $\alpha = 1.2$, a random single-node attack leads to more failures than a targeted attack on the node with the largest degree. This is because that V_1 becomes overloaded *almost surely* with $\alpha \to 1$. Thus, when a node with a larger degree is attacked, more failures will be triggered. However, when α becomes larger, the random single-node attack brings more failures. This is the result of Lemma 3.1.1, whereby failure condition (3.2) is less likely to be satisfied when the initial-attacked node has a larger degree with a large α. Besides, it is shown that the failure propagation under these three attack methods reaches to steady state at the almost same time. Therefore, our analysis results can be applied as an estimate of steady state for these two attack methods.

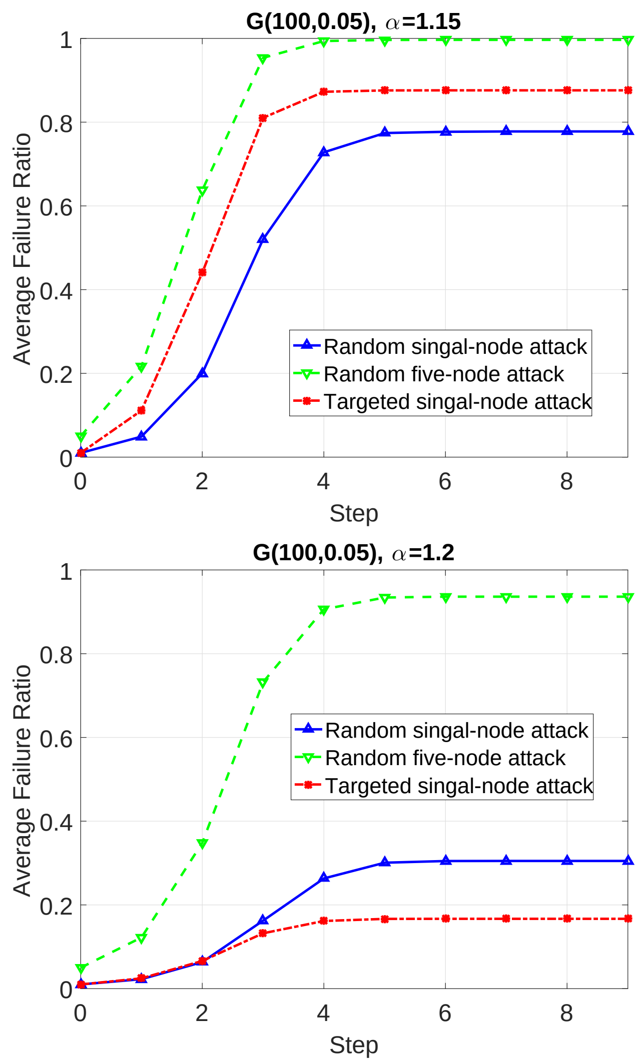

Figure 3.4: Step-by-step comparison of the average failure ratios under three attack methods in $G(100, 0.05)$. Under a targeted single-node attack, the node with largest degree is attacked. A random five-node attack is always more devastating than a random single-node attack. A targeted single-node attack leads to more failures than a random single-node attack with $\alpha = 1.15$ but fewer failures than a random single-node attack with $\alpha = 1.2$. Reprinted from [11].

4. PHASE TRANSITION OVER α^*

As we mentioned earlier, the average failure ratio decreases from near 1 to $1/n$ as α varies over the interval $(1, 2)$. Simulation results in Fig. 3.1, however, indicate a rapid phase transition over a much smaller interval, denoted in this paper by the *threshold interval*. In this section, we study this interval to further characterize the network reaction to random single-node attacks. We first find a *critical* point in this interval at which the failure ratio becomes very sensitive to the changes in α. We then derive the lower and upper bounds of the interval.

4.1 Critical Value of α

For each $E[f_t],\ t \geq 1$, we define a critical value of α at which $E[f_t]$ takes the median value of its variable range. As seen in Fig. 3.1, $E[f_t]$ undergoes rapid changes in the close vicinity of such a critical value. Since

$$E[f_t] = \frac{1}{n} E[| \cup_{d=1}^{t} V_d|] \cdot \Pr\{v \text{ fails} \mid v \in \cup_{d=1}^{t} V_d\} + \frac{1}{n}, \tag{4.1}$$

where $\Pr\{v \text{ fails} \mid v \in \cup_{d=1}^{t} V_d\}$ has a range $[0, 1]$, the variable range of $E[f_t]$ is

$$E[f_t] \in \left[\frac{1}{n}, \frac{E[| \cup_{d=1}^{t} V_d|] + 1}{n}\right]. \tag{4.2}$$

*Reprinted from [11] "Load-Dependent Cascading Failures in Finite-Size Erdös-Rényi Random Networks" by D. Lv and A. Eslami and S. Cui, 2017, Network Science and Engineering, IEEE Trans on (TNSE), Volume 4, Page 129-139. Copyright 2017 by IEEE.

According to (4.1), the critical value of α could be found when $\Pr\{v \text{ fails} \mid v \in \cup_{d=1}^{t} V_d\} = 0.5$. Let $\alpha_{c,t}$ denote the critical α at step t, which is formally defined as

$$\alpha_{c,t} \triangleq \arg_{\alpha}\left\{\Pr\{v \text{ fails} \mid v \in \cup_{d=1}^{t} V_d\} = \frac{1}{2}\right\}, \ \forall t \geq 1. \tag{4.3}$$

We start our analysis by finding $\alpha_{c,1}$, i.e., the critical α for $E[f_1]$. In the following, we drop the index t for $\alpha_{c,t}$, $\alpha_{l,t}$ and $\alpha_{h,t}$, wherever it is clear from the text. The proof of Theorem 4.1.1, including the justification of the assumption is provided in the appendix.

Theorem 4.1.1 Consider a random single-node attack applied to $G(n, p)$. Under the same assumption of Theorem 1. The critical value of α for the first step, i.e., $\alpha_{c,1}$, can be obtained by

$$\alpha_{c,1} \approx 1 + \frac{1}{1 + (n-2)p}.$$

Before deriving $\alpha_{c,t}$ for $t \geq 2$, consider Fig. 4.1, which shows the average failure ratio over $G(100, 0.05)$ via simulations. The four curves shown are recorded at steps 1, 2, 3, and 9 of the cascade. These curves suggest that the variations of $\alpha_{c,t}$ over different steps are very small. In fact, in $G(100, 0.05)$, $\alpha_{c,t}$ for $t \geq 1$ can be numerically obtained as 1.17 ± 0.05. The following set of analytical results shed light on the fact that the critical α stays almost the same across different steps, suggesting that $\alpha_{c,1}$ may very well be used to approximate $\alpha_{c,t}$, $t \geq 2$.

Lemma 4.1.1 Consider a random single-node attack applied to $G(n, p)$. At any arbitrary step, the failure probability of any node or set of nodes is a non-increasing function of α.

proof 3 This lemma is intuitive. If node e fails with $\alpha = \alpha^*$, it also fails if $\alpha \in (1, \alpha^*)$. The same is true for the failures of sets of nodes. Corollary 4.1.1 can be proven similarly.

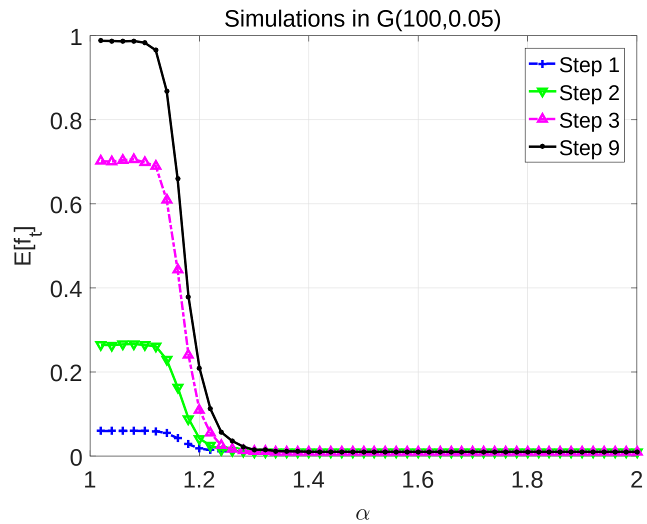

Figure 4.1: Comparison of critical values of α at different steps. Reprinted from [11].

Corollary 4.1.1 Consider a random single-node attack applied to $G(n,p)$. At any arbitrary step of the failure propagation, the failure probabilities $\Pr\{V_1 \textit{ fails}\}$, $\Pr\{V_1 \cup V_2 \textit{ fail}\}, \cdots, \Pr\{\bigcup_d V_d \textit{ fail}\}$ are non-increasing functions of α.

Notice that no assumptions are required for Lemma 4.1.1 and Corollary 4.1.1.

Theorem 4.1.2 Consider a random single-node attack applied to $G(n,p)$ with $\alpha \in (1, \frac{1}{(n-2)p} + 1]$, and define a family of events: X_t="V_t fails at step t", $t \geq 1$. We assume that the set F_{t-1} can be considered as a large virtual node that redistributes its load to its alive neighbors at step t. Then the conditional probabilities of failure satisfy $\Pr\{X_t \mid X_{t-1}\} \geq \Pr\{X_1\}, \forall t > 1$.

The proof of Theorem 4.1.2, including the justifications of the assumptions used, is provided in the appendix. First note that Lemma 4.1.1 along with Theorem 4.1.1 indicates that $\Pr\{V_1 \text{ fails at step } 1\} \leq 0.5$ when $\alpha > \alpha_{c,1} \approx \frac{1}{1+(n-2)p} + 1$. Since any further failures are conditioned on the failure of V_1 in the first step, the failure probability of any node at any step will be smaller than or equal to $\Pr\{V_1 \text{ fails at step } 1\} \leq 0.5$ when

$\alpha > \alpha_{c,1} \approx \frac{1}{1+(n-2)p} + 1$. This along with the definition of $\alpha_{c,t}$ implies that $\alpha_{c,t}$ for every step $t \geq 2$ is within the small range $(1, \alpha_{c,1})$. Note that $p = \ln n/n$ is the threshold for connectedness of $G(n,p)$ [5]. Therefore, in a connected $G(n,p)$, we have $np \geq \ln n$, which is greater than or equal to 3.22 in typical-size networks of 25 nodes or larger. This leads to $\alpha_{c,1}$ is approximate less or equal than 1.24 for connected networks that are of interest in this paper. On the other hand, we can see from Fig. 4.1 that there exists an $\alpha_l > 1$ (to be calculated in Section 4.2) such that $\Pr\{V_1 \text{ fails at step } 1\} \approx 1$ for $\alpha \in (1, \alpha_l]$. By applying Theorem 4.1.2 and the chain rule of probability, we obtain $\Pr\{\cup_{d=1}^{t} V_d \text{ fail}\} \geq (\Pr\{V_1 \text{ fails at step } 1\})^t \approx 1$ for $\alpha \in (1, \alpha_l]$. Therefore, $\alpha_{c,t}$ sits within a very narrow range of $(\alpha_l, \alpha_{c,1})$, enabling us to approximate $\alpha_{c,t},\ t \geq 2$ by $\alpha_{c,1}$ given in Theorem 4.1.1.

Numerical results are shown in Table 4.1 to quantify the accuracy of approximating $\alpha_{c,t}, t \geq 2$ by $\alpha_{c,1}$. Relative approximation errors are provided for various steps of cascade over four different topologies: $G(100, 0.03)$, $G(100, 0.05)$, $G(200, 0.03)$ and $G(400, 0.01)$. Under each topology, $\alpha_{c,t}, t = 1, \cdots, 9$, are first manually obtained from simulations according to the definition of $\alpha_{c,t}$, and then compared against $\alpha_{c,1}$ given by Theorem 4.1.1. As we see, the approximation error in each case is less than 1%, suggesting that the simple expression for $\alpha_{c,1}$ given by Theorem 4.1.1 can be applied to approximate the critical value of the tolerance parameter for almost every step.

Because that the close-form expression of $\alpha_{c,1}$ given by Theorem 4.1.1 is simple and only depend on np and p, not depend on n. This result can even be applied to the asymptotic case: $n \to \infty, np \to \lambda$. In such case, we have $p \to 0$ and $\alpha_c \to 1 + 1/(1+\lambda)$. This knowledge of α_c can be utilized as a fast prediction of the α region where the cascading failures happen in a network.

4.2 Threshold Interval $[\alpha_l, \alpha_h]$

As we discussed at the beginning of this section, $E[f_t]$ drops significantly from its maximum value to near zero within a small interval, referred to as the threshold interval. Based on the relationship between $E[f_t]$ and $\Pr\{v \text{ fails} \mid v \in \cup_{d=1}^{t} V_d\}$ given in (4.1), this is the interval where $\Pr\{v \text{ fails} \mid v \in \cup_{d=1}^{t} V_d\}$ drops from near 1 to $1/n$. We denote the threshold interval at step t by $[\alpha_{l,t}, \alpha_{h,t}]$. With $\alpha_{l,t}$ and $\alpha_{h,t}$, the domain of α at step t is divided into three parts: $(1, \alpha_{l,t}), [\alpha_{l,t}, \alpha_{h,t}]$ and $(\alpha_{h,t}, 2)$. In particular, $\Pr\{v \text{ fails} \mid v \in \cup_{d=1}^{t} V_d\}$ stays very close to 1 for $\alpha \in (1, \alpha_{l,t})$, drops quickly for $\alpha \in [\alpha_{l,t}, \alpha_{h,t}]$, and stays close to 0 for $\alpha \in (\alpha_{h,t}, 2)$. In this section, we provide approximations for $\alpha_{l,t}$ and $\alpha_{h,t}$ and verify their accuracy via simulations.

As a well-studied result, the normal random variable takes 95% and 99% of its values within two and three standard deviations of its mean, respectively. This fact is known as the *empirical rule*. Accordingly, we present two criteria for the threshold interval, namely 2σ and 3σ intervals, within which $\Pr\{v \text{ fails} \mid v \in \cup_{d=1}^{t} V_d\}$ loses 95% and 99% of its maximum values, respectively. We start by presenting our method to approximate the threshold interval for the first step of a cascade. The proof of Theorem 4.2.1, including the justifications of the assumptions, is provided in the appendix.

Theorem 4.2.1 Consider a random single-node attack applied to $G(n,p)$. Under the same assumption of Theorem 1. We also assume $\alpha_{h,1}$ and $\alpha_{l,1}$ are estimated given $|V_1| = (n-1)p$. The threshold interval $[\alpha_{l,1}, \alpha_{h,1}]$ for the first step of the cascade can be obtained as follows:

$$\begin{aligned}
\alpha_{h,1} &\approx 1 + \frac{1}{-c\sqrt{(p + \frac{n-3}{n-1})(1-p)} + 1 + (n-2)p}, \\
\alpha_{l,1} &\approx 1 + \frac{1}{c\sqrt{(p + \frac{n-3}{n-1})(1-p)} + 1 + (n-2)p},
\end{aligned}$$

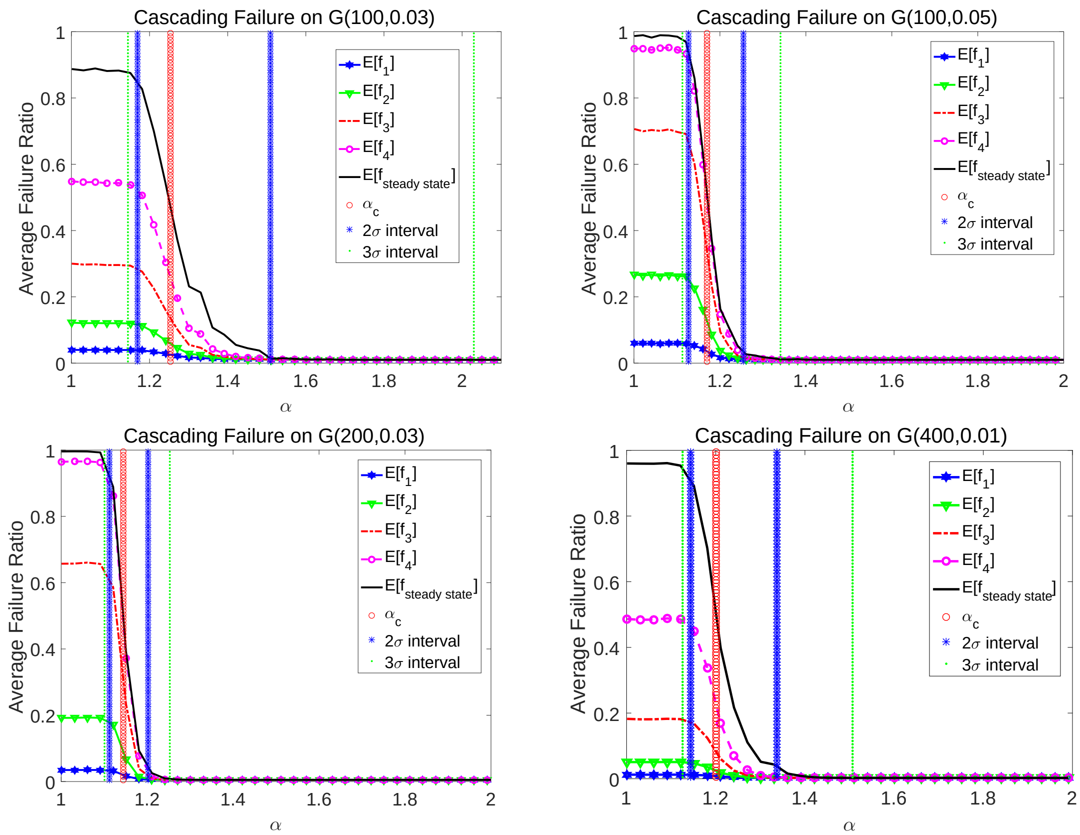

Figure 4.2: Verification of α_c and threshold intervals for $G(100, 0.03)$, $G(100, 0.05)$, $G(200, 0.03)$, $G(400, 0.01)$; $E[f_1]$, $E[f_2]$, $E[f_3]$, $E[f_4]$, $E[f_{\text{steady state}}]$ were obtained from cascading failure simulations; α_c and 2σ interval, 3σ interval were given by Theorem 4.1.1 and Theorem 4.2.1. Reprinted from [11].

where $c = 2$ and $c = 3$ lead to the 2σ and 3σ threshold intervals, respectively. Within the 2σ and 3σ threshold intervals, $E[f_1]$ drops 95% and 99% of its maximum value, respectively.

In Section 4.1, we showed how $\alpha_{c,t}, t \geq 2$ is well approximated by $\alpha_{c,1}$. Similarly, we justify here how $\alpha_{l,1}$ and $\alpha_{h,1}$ above serve as good approximations of $\alpha_{l,t}$ and $\alpha_{h,t}$, $t \geq 2$. Since all future failures are conditioned on the failure of the first step, we have $\Pr\{v \text{ fails} \mid v \in \cup_{d=1}^{t} V_d\} \leq \Pr\{V_1 \text{ fails at step } 1\}, \forall \alpha \in (1, 2)$. Therefore, the definition of threshold interval yields

$$\alpha_{l,t} \leq \alpha_{l,1},\ \alpha_{h,t} \leq \alpha_{h,1}. \tag{4.4}$$

However, applying Theorem 4.1.2 and the chain rule of probability to the interval $\alpha \in (1, \alpha_{l,1}]$, we have

$$\begin{aligned} &\Pr\{v \text{ fails} \mid v \in \cup_{d=1}^{t} V_d\} \geq \Pr\{\cup_{d=1}^{t} V_d \text{ fail}\} \\ &\geq (\Pr\{V_1 \text{ fails at step } 1\})^t,\ \alpha \in (1, \alpha_{l,1}]. \end{aligned} \tag{4.5}$$

Since $\Pr\{V_1 \text{ fails at step } 1\} \approx 1$ when $\alpha \in (1, \alpha_{l,1}]$, we obtain

$$\begin{aligned} \Pr\{v \text{ fails} \mid v \in \cup_{d=1}^{t} V_d\} \ &\approx 1,\ \alpha \in (1, \alpha_{l,1}] \\ &\Rightarrow \alpha_{l,t} \geq \alpha_{l,1}. \end{aligned} \tag{4.6}$$

From (4.4) and (4.6) we conclude that $\alpha_{l,t} \approx \alpha_{l,1}$, $\forall t \geq 2$. On the other hand, we know that $\alpha_{h,t} \in (\alpha_{c,t}, \alpha_{h,1}]$, which is very narrow. Therefore, $\alpha_{h,1}$ provides a tight upper bound for $\alpha_{h,t}$, which is often appreciated when designing a robust network. Here, we simply approximate $\alpha_{h,t}$ by $\alpha_{h,1}$ given in Theorem 4.2.1.

We now verify the accuracy of our findings for α_c and $[\alpha_l, \alpha_h]$ by considering cascading failures over four different topologies: $G(100, 0.03)$, $G(100, 0.05)$, $G(200, 0.03)$ and $G(400, 0.01)$. The exact values obtained by simulations are depicted in Fig. 4.2, along with the analytical approximates given by Theorems 4.1.1 and 4.2.1. Simulation results were obtained in the same method with Fig. 3.1. As it can be seen, α_c and $[\alpha_l, \alpha_h]$ obtained from Theorem 4.1.1 and Theorem 4.2.1 accurately characterize the phase transition in the average failure ratio due to a random single-node attack. This knowledge of threshold interval can be utilized to avoid risky or redundant investment when the tolerance parameter is set to operate the network.

Table 4.1: True values of $\alpha_{c,t}$ for steps 1 through 9, their approximations by $\alpha_{c,1}$ from Theorem 4.1.1, and the approximation errors. Reprinted from [11].

$n = 100, p = 0.03$

Step t	$\alpha_{c,t}$	$\alpha_{c,1}$ by Theorem 3	Relative Error
1	1.2530		0.0644%
2	1.2551		0.1030%
3	1.2521		0.1363%
4	1.2521	1.2538	0.1363%
5	1.2551		0.1030%
6	1.2551		0.1030%
9	1.2551		0.1030%

$n = 100, p = 0.05$

Step t	$\alpha_{c,t}$	$\alpha_{c,1}$ by Theorem 3	Relative Error
1	1.1750		0.4688%
2	1.1711		0.1373%
3	1.1711		0.1373%
4	1.1711	1.1695	0.1373%
5	1.1741		0.3925%
6	1.1741		0.3925%
9	1.1741		0.3925%

$n = 200, p = 0.03$

Step t	$\alpha_{c,t}$	$\alpha_{c,1}$ by Theorem 3	Relative Error
1	1.1531		0.7812%
2	1.1441		0%
3	1.1471		0.2622%
4	1.1471	1.1441	0.2622%
5	1.1471		0.2622%
6	1.1471		0.2622%
9	1.1471		0.2622%

$n = 400, p = 0.01$

Step t	$\alpha_{c,t}$	$\alpha_{c,1}$ by Theorem 3	Relative Error
1	1.2000		0.0669%
2	1.1991		0.1418%
3	1.1940		0.5999%
4	1.1933	1.2008	0.6248%
5	1.1999		0.0752%
6	1.1200		0.0669%
9	1.1200		0.0669%

5. CONCLUSION*

We introduced a load-based cascade model to study the vulnerability of complex networks under random single-node attacks, where the ER random graph with finite size was used to represent the network. We assumed that the capacity of a node is proportional to its initial load and the load of a failed node is redistributed to its neighbors according to their capacity. The average failure ratio at each step was used to quantify the damage experienced by the network. A step-by-step estimation of the average failure ratio has been provided. The accuracy of such estimations was validated by numerical results. Our analysis for finite-size networks revealed a phase transition phenomenon in network reactions to single-node attacks, where the average value of the failure ratio drops quickly within a short interval of the load margin. We characterized this interval by finding the critical value of the tolerance parameter at which the failure ratio takes its median value and is most sensitive to the variation of the tolerance parameter. We also derived the threshold interval within which this phase transition occurs. Our findings shed light on how to set the load margin for both robustness and efficient use of resources in designing networks resilient to random single-node attacks.

*Reprinted from [11] "Load-Dependent Cascading Failures in Finite-Size Erdös-Rényi Random Networks" by D. Lv and A. Eslami and S. Cui, 2017, Network Science and Engineering, IEEE Trans on (TNSE), Volume 4, Page 129-139. Copyright 2017 by IEEE.

REFERENCES

[1] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang. Complex networks: Structure and dynamics. *Physics Reports*, 424(45):175–308, Jan. 2006.

[2] Paolo Crucitti, Vito Latora, and Massimo Marchiori. Model for cascading failures in complex networks. *Phys. Rev. E*, 69(4):045104, Apr. 2004.

[3] I. Dobson, B.A. Carreras, and D.E. Newman. A branching process approximation to cascading load-dependent system failure. In *Proceedings of the 37th Annual Hawaii International Conference on System Sciences*, pages 1–10, Hawaii, Jan. 2004.

[4] P. Erdös and A. Rényi. On random graphs I. *Publ. Math. Debrecen*, 6:290–297, 1959.

[5] Paul Erdös and A Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci*, 5:17–61, 1960.

[6] E. N. Gilbert. Random graphs. *The Annals of Mathematical Statistics*, 30(4):1141–1144, 1959.

[7] Ake J. Holmgren. Using graph models to analyze the vulnerability of electric power networks. *Risk Analysis*, 26(4):955–969, Aug. 2006.

[8] R. Kinney, P. Crucitti, R. Albert, and V. Latora. Modeling cascading failures in the North American power grid. *The European Physical Journal B - Condensed Matter and Complex Systems*, 46(1):101–107, July 2005.

[9] Jure Leskovec and Andrej Krevl. SNAP Datasets: Stanford large network dataset collection. http://snap.stanford.edu/data, June 2014.

[10] Daqing Li, Bowen Fu, Yunpeng Wang, Guangquan Lu, Yehiel Berezin, H. Eugene Stanley, and Shlomo Havlin. Percolation transition in dynamical traffic network with

evolving critical bottlenecks. *Proceedings of the National Academy of Sciences*, 112(3):669–672, 2015.

[11] D. Lv, A. Eslami, and S. Cui. Load-dependent cascading failures in finite-size erdos-renyi random networks. *IEEE Transactions on Network Science and Engineering*, 4(2):129–139, April 2017.

[12] Yi-Hua Ma and Dong-Li Zhang. Cascading network failure based on local load distribution and non-linear relationship between initial load and capacity. In *Machine Learning and Cybernetics (ICMLC), 2012 International Conference on*, volume 3, pages 935–940, July 2012.

[13] Adilson E. Motter and Ying-Cheng Lai. Cascade-based attacks on complex networks. *Phys. Rev. E*, 66(6):065102, Dec. 2002.

[14] M. E. J. Newman, S. H. Strogatz, and D. J. Watts. Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E*, 64:026118, Jul 2001.

[15] Mark EJ Newman. Random graphs as models of networks. *arXiv preprint cond-mat/0202208*, Feb 2002.

[16] Ke Sun and Zhen-Xiang Han. Analysis and comparison on several kinds of models of cascading failure in power system. In *Transmission and Distribution Conference and Exhibition: Asia and Pacific, IEEE/PES*, pages 1–7, Dalian, China, Aug. 2005.

[17] JianWei Wang and LiLi Rong. Cascade-based attack vulnerability on the US power grid. *Safety Science*, 47(10):1332–1336, Dec. 2009.

[18] JianWei Wang and LiLi Rong. Robustness of the western United States power grid under edge attack strategies due to cascading failures. *Safety Science*, 49(6):807–812, July 2011.

[19] Duncan J. Watts. A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences*, 99(9):5766–5771, Apr. 2002.

[20] ZhiXi Wu, Gang Peng, WenXu Wang, Sammy Chan, and Eric WingMing Wong. Cascading failure spreading on weighted heterogeneous networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(05):P05013, May 2008.

APPENDIX A

SOME PROOFS FOR THEOREMS, COROLLARIES, LEMMAS

Lemma 2.3.1 Consider a random single-node attack applied to $G(n, p)$. Let node a be the attacked node, and node e be an arbitrary node in $V \backslash \{a\}$. Let P_d be the probability that the shortest path from e to a has length d; $\Pr\{B_d\}$ be the probability that at least one path from e directly through a node in $V \backslash \{a \cup e\}$ to a has a length which is less than or equal to d. Then $E[|V_d|], d \geq 1$, the average size of V_d, is simply

$$E[|V_d|] \;=\; (n-1)P_d,$$

where $P_d, d \geq 1$, can be obtained recursively as

$$\begin{aligned} P_1 &= p, \\ P_2 &= (1-p)(1-(1-p^2)^{n-2}), \\ P_d &= (1-p)\Pr\{B_d\} - \sum_{j=2}^{d-1} P_j,\ d > 2. \end{aligned}$$

In the numerical calculation, we assume that $\Pr\{B_d\}, d > 2$, can be approximated recursively as

$$\Pr\{B_d\} \;\approx\; 1 - ((1-p) + p \cdot (1 - \sum_{j=1}^{d-1} P_j))^{n-2}.$$

proof 4 (Proof of Lemma 2.3.1) Suppose that we randomly pick a node a from the ER graph $G(n, p)$, where any two nodes are connected with probability p. For an arbitrary node $e \in V \backslash a$, we define a family of probabilities as $P_d = \Pr\{$the shortest path between *nodes e and a has length d}, $P_{[i,j]} = \Pr\{$the shortest path between nodes e and a has length within $[i,j]\}$, satisfying*

$$P_{[i,j]} = \sum_{d=i}^{j} P_d, \quad 0 < i \leq j.$$

Since P_d is the same for all the nodes in $V \backslash a$, the expectation of $|V_d|$ over all possible topologies can be calculated as

$$E[|V_d|] = (n-1)P_d. \tag{5.1}$$

We use induction to obtain P_d. We first have $P_1 = p$. When $d \geq 2$, we find P_d as

$$P_d = \begin{cases} P_{[2,d]} & d = 2 \\ P_{[2,d]} - \sum_{j=2}^{d-1} P_j & d > 2 \end{cases}, \tag{5.2}$$

where $P_{[2,d]}$ remains to be found. The event "the shortest distance from e to a is within $[2,d]$" is true if the following two independent events happen at the same time: $A =$"e is not directly connected to a" and $B_d =$"at least one path from e directly through a node in $V \backslash \{a \cup e\}$ to a has a length which is less than or equal to d". It can be easily seen that $\Pr\{A\} = 1 - p$. Then $P_{[2,d]},\ d \geq 2$ can be obtained as

$$\begin{aligned} P_{[2,d]} &= \Pr\{A\} \Pr\{B_d\} \\ \xrightarrow{\Pr\{A\}=1-p} &= (1-p) \Pr\{B_d\}. \end{aligned} \tag{5.3}$$

Now we aim to obtain $\Pr\{B_d\},\ d \geq 2$. Given $d = 2$, node e can go through any node in $V \backslash \{a \cup e\}$ to a with probability p^2. These $(n-2)$ possible paths are independent. B_2

happens if any of these paths is connected. Therefore,

$$\Pr\{B_2\} \;=\; 1-(1-p^2)^{n-2}. \tag{5.4}$$

Combining (5.2)*,* (5.3) *and* (5.4) *we have*

$$P_2 \;=\; (1-p)(1-(1-p^2)^{n-2}). \tag{5.5}$$

When $d > 2$, substituting (5.3) *into* (5.2)*, we have*

$$P_d \;=\; (1-p)\Pr\{B_d\} - \sum_{j=2}^{d-1} P_j,\; d > 2. \tag{5.6}$$

Combining $P_1 = p$, (5.5) *and* (5.6) *yields Lemma 1.*

Now we aim to verify the approximation

$$\Pr\{B_d\} \;\approx\; 1-((1-p)+p\cdot(1-\sum_{j=1}^{d-1} P_j))^{n-2}, d > 2. \tag{5.7}$$

Since directly obtaining $\Pr\{B_d\}$ is complicated, we first focus on its complement $\bar{B}_d =$"the lengths of the $(n-2)$ shortest paths from e directly through a node in $V\backslash\{a\cup e\}$ to a are all greater than d". Therefore, $\Pr\{B_d\}$ can be obtained through $1-\Pr\{\bar{B}_d\}$. Now our goal is to approximate $\Pr\{\bar{B}_d\}$. For notational simplicity, let $l_1, \cdots, l_{n-2}$ denote the lengths of the $(n-2)$ shortest paths from e directly through a node in $V\backslash\{a\cup e\}$ to a, such that $\Pr\{\bar{B}_d\}$ is the joint probability that $l_1, \cdots, l_{n-2}$ are greater than d, which can be expressed as

$$\Pr\{\bar{B}_d\} = \Pr\{\bigcap_{i=1}^{n-2} l_i > d\}. \tag{5.8}$$

To obtain (5.8)*, we start from analyzing the joint probability that arbitrary two path lengths from* $l_1, \cdots, l_{n-2}$ *are greater than* d*. Let us randomly pick two nodes* $v_1, v_2 \in V \backslash \{a \cup e\}$ *and consider paths* $e \rightarrow v_1 \dashrightarrow a$ *and* $e \rightarrow v_2 \dashrightarrow a$*, where "*$\dashrightarrow$*" stands for the shortest path between the two nodes. Let* l_1 *denote the length of* $e \rightarrow v_1 \dashrightarrow a$ *and* l_2 *denote the length of* $e \rightarrow v_2 \dashrightarrow a$*. In ER random graph,* e *is connected to each node independently with probability* p*. We can list the following three scenarios regarding the connectivity between* e *and* v_1, v_2*:*

1. e *is not directly connected to* v_1*, such that* $l_1 = \infty$*. In this case,* l_1 *and* l_2 *will not be affected by each other, such that they are independent.*

2. e *is not directly connected to* v_2*. Similarly,* l_1 *and* l_2 *are independent.*

3. e *is directly connected to both* v_1 *and* v_2*. In this case* l_1 *and* l_2 *become dependent.*

Let $C_1 =$*"Scenario 3):* e *is directly connected to both* v_1 *and* v_2*". According to scenarios 1) and 2), we have* $l_1 \perp l_2 \mid \bar{C}_1$*. Now we aim to calculate the joint probability* $\Pr\{l_1 > d, l_2 > d\}$*. Given* $d > 2$*,* $\Pr\{l_1 > d, l_2 > d\}$ *can be rewritten according to the law of total probability:*

$$
\begin{aligned}
& \Pr\{l_1 > d, l_2 > d\} \\
= \; & \Pr\{l_1 > d, l_2 > d \mid C_1\} \Pr\{C_1\} + \\
& \Pr\{l_1 > d, l_2 > d \mid \bar{C}_1\} \Pr\{\bar{C}_1\} \\
= \; & \Pr\{l_1 > d, l_2 > d \mid C_1\} p^2 + \\
& \Pr\{l_1 > d, l_2 > d \mid \bar{C}_1\} (1 - p^2) \\
\xrightarrow{l_1 \perp l_2 \mid \bar{C}_1} = \; & \Pr\{l_1 > d, l_2 > d \mid C_1\} p^2 + \\
& \Pr\{l_1 > d \mid \bar{C}_1\} \Pr\{l_2 > d \mid \bar{C}_1\} (1 - p^2),
\end{aligned}
\tag{5.9}
$$

where $\Pr\{l_1 > d \mid \bar{C}_1\}$ *can be obtained by*

$$\begin{aligned}\Pr\{l_1 > d \mid \bar{C}_1\} &= (\Pr\{l_1 > d\} - \Pr\{l_1 > d \mid C_1\} \Pr\{C_1\}) / \Pr\{\bar{C}_1\} \\ &= (\Pr\{l_1 > d\} - \Pr\{l_1 > d \mid C_1\} p^2)/(1 - p^2). \end{aligned} \tag{5.10}$$

Similar to (5.10)*, we can obtain* $\Pr\{l_2 > d \mid \bar{C}_1\}$*. Substituting them into* (5.9)*, we obtain*

$$\begin{aligned}\Pr\{l_1 > d, l_2 > d\} &= \Pr\{l_1 > d, l_2 > d \mid C_1\} p^2 + \\ &\quad (\Pr\{l_1 > d\} - \Pr\{l_1 > d \mid C_1\} p^2) \\ &\quad \cdot (\Pr\{l_2 > d\} - \Pr\{l_2 > d \mid C_1\} p^2)/(1 - p^2), \\ &\xrightarrow{\textit{When } p^2 \approx 0, \textit{ cancel } p^2} \approx \Pr\{l_1 > d\} \Pr\{l_2 > d\}. \end{aligned} \tag{5.11}$$

The approximation (5.11) *holds when* p^2 *is small, which is true under typical values of* p *in networks of typical sizes, e.g.,* $n \geq 20$*. In addition, some simulations were conducted to test the accuracy of approximation* (5.11)*, and the results support our analysis. The following steps describe how the simulations were conducted: in a network of size* n*, arbitrarily four nodes* a, e, v_1*, and* v_2 *were selected. Then we counted* l_1 *and* l_2 *in* $100,000$ *realizations of ER random graph. Based on the numerical results of* l_1 *and* l_2*, we calculated* $\Pr\{l_1 > d, l_2 > d\}$ *and* $\Pr\{l_1 > d\} \Pr\{l_2 > d\}$*. Under varied values of* n*,* p *and* d*, these two probabilities are always approximately equal, which indicates the assumption is valid even in small networks with relatively larger* p*, e.g.,* $n = 20, p = 0.3$*. Partial results are shown in Table 5.1.*

Based on (5.11)*, we aim to extend the approximation to the joint probability* $\Pr\{\bar{B}_d\}$

in (5.8)*. First we rewrite* (5.11) *as*

$$\begin{aligned}
&\Pr\{l_1 > d, l_2 > d\} \\
=\ &\Pr\{l_1 > d \mid l_2 > d\} \Pr\{l_2 > d\} \\
\approx\ &\Pr\{l_1 > d\} \Pr\{l_2 > d\}, \\
&\xrightarrow{\textit{Since } \Pr\{l_2>d\}\neq 0,\ \textit{cancel } \Pr\{l_2>d\} \textit{ on both sides}} \\
&\Pr\{l_1 > d \mid l_2 > d\} \ \approx\ \Pr\{l_1 > d\}. \qquad (5.12)
\end{aligned}$$

Since an ER random graph is homogeneous network and each node has an identical statistical property, result (5.12) *also applies to other path lengths* $l_1, \cdots, l_{n-2}$*. In a finite-size network* $G(n,p)$*, according to the chain rule, the joint probability* $\Pr\{\bar{B}_d\} = \Pr\{\bigcap_{i=1}^{n-2} l_i > d\}$ *can be expanded as*

$$\begin{aligned}
\Pr\{\bar{B}_d\} \ &= \Pr\{\bigcap_{i=1}^{n-2} l_i > d\} \\
&= \prod_{i=1}^{n-2} \Pr\{l_i > d \mid \bigcap_{j=1}^{i-1} l_j > d\}, \\
\xrightarrow{(5.12)} &\approx \prod_{i=1}^{n-2} \Pr\{l_i > d\}. \qquad (5.13)
\end{aligned}$$

In order to find (5.13)*, we need to have* $\Pr\{l_i > d\}, i = 1, \cdots, n-2$*. Assume* l_1 *is the length of the path through node* v_1*, so* $l_1 > d$ *happens when either* e *is not connected to* v_1 *(*$l_1 = \infty$*) or* e *is connected to* v_1 *but the distance between* v_1 *and* a *is greater than* $d-1$*. Because ER random graph is a homogeneous network and* $l_1, \cdots, l_{n-2}$ *have the identical statistical property,* $\Pr\{l_i > d\},\ i = 1, 2, \cdots, n-2$*, can be obtained by*

$$\Pr\{l_i > d\} \ = \ (1-p) + p \cdot (1 - P_{[1,d-1]}). \qquad (5.14)$$

Substituting (5.14) *into* (5.13)*, we have*

$$\Pr\{\bar{B}_d\} \approx ((1-p) + p \cdot (1 - \sum_{j=1}^{d-1} P_j))^{n-2}, \ d > 2, \tag{5.15}$$

Thus, the approximation in (5.7) *is justified.*

Table 5.1: Simulation results to test the dependence of two path lengths.

Network	d	$\Pr\{l_1 > d, l_2 > d\}$	$\Pr\{l_1 > d\} \cdot \Pr\{l_2 > d\}$
$G(100, 0.05)$	4	0.9287	0.9286
	10	0.9054	0.9051
$G(20, 0.3)$	4	0.8265	0.8270
	8	0.4926	0.4922

Theorem 3.2.1 Consider a random single-node attack applied to $G(n,p)$. We assume the conditional distribution of $k(V_1)$ given $|V_1| = x$ is approximately normal with mean μ and variance σ^2, where

$$\begin{aligned} \mu &= x + x(x-1)p + (n-x-1)xp, \\ \sigma^2 &= (2x(x-1) + (n-x-1)x)p(1-p). \end{aligned}$$

Then $E[f_1]$, i.e., the average failure ratio at step 1, can be approximated as

$$E[f_1] \approx \frac{1}{n}\left(1 + \sum_{x=1}^{n-1} x \cdot \binom{n-1}{x} p^x (1-p)^{n-1-x} \Phi(\frac{\frac{x}{\alpha-1} - \mu}{\sigma})\right),$$

where $\binom{n-1}{x} = \frac{(n-1)!}{x!(n-x-1)!}$ is the binomial coefficient, and $\Phi(\cdot)$ is the cumulative distribution function (CDF) for the standard normal distribution.

proof 5 (Proof of Theorem 3.2.1) Since $E[f_1]$ can be obtained as $E[f_1] = \frac{1}{n}(1 + E[|F_1|])$, it is enough to find $E[|F_1|]$. Based on Corollary 1, we have the failure condition for V_1 as $k(V_1) < \frac{k(V_0)}{\alpha-1}$. According to the failure condition, the distribution of $|F_1|$ can be expressed as

$$\begin{aligned} &\Pr\{|F_1| = x\} \\ = &\Pr\{k(V_1) < \frac{k(V_0)}{\alpha - 1} \mid |V_1| = x\} \cdot \Pr\{|V_1| = x\}, \end{aligned} \tag{5.16}$$

and expectation of $|F_1|$ can be obtained by the law of total probability as

$$\begin{aligned} &E[|F_1|] = \\ &\sum_{x=1}^{n-1} x \cdot \Pr\{k(V_1) < \frac{k(V_0)}{\alpha - 1} \mid |V_1| = x\} \cdot \Pr\{|V_1| = x\}, \end{aligned} \tag{5.17}$$

where $|V_1|$ is the number of nodes in V_1, which obeys a binomial distribution. That is, $|V_1| \sim B(n-1, p)$ and

$$\Pr\{|V_1| = x\} = \binom{n-1}{x} p^x (1-p)^{n-1-x}. \tag{5.18}$$

Now to find (5.17), we need to calculate the conditional distribution of $k(V_1)$ given $|V_1| = x$. The links adjacent to nodes in V_1 can be divided into three categories: edges between V_0 and V_1, within V_1, and between V_1 and V_2, denoted by the sets $\mathcal{E}(V_0, V_1)$, $\mathcal{E}(V_1)$, and $\mathcal{E}(V_1, V_2)$, respectively. Such partition of edges is illustrated in Fig. 5.1. We then have

$$k(V_1) \ = \ |\mathcal{E}(V_0, V_1)| + 2|\mathcal{E}(V_1)| + |\mathcal{E}(V_1, V_2)|. \tag{5.19}$$

Here we show $k(V_1)$ is approximately normal given $|V_1| = x,\ x \in [1, n-1]$. In (5.19), it

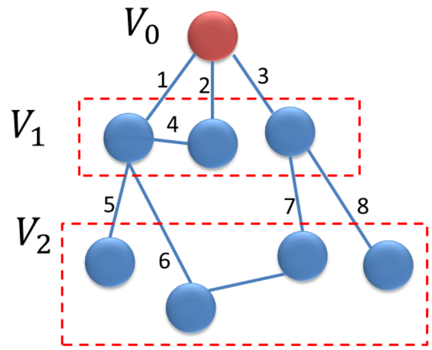

Figure 5.1: Partition of edges within and adjacent to V_1: $\mathcal{E}(V_0, V_1) = \{1, 2, 3\}, \mathcal{E}(V_1) = \{4\}, \mathcal{E}(V_1, V_2) = \{5, 6, 7, 8\}$.

can be easily seen that $|\mathcal{E}(V_0, V_1)| = x$, while $|\mathcal{E}(V_1)|$ and $|\mathcal{E}(V_1, V_2)|$ depend on connectivity of nodes. In a ER random graph, each pair of nodes are connected with a probability p independent of other pairs [6]. Thus, $|\mathcal{E}(V_1)|$ follows a binomial distribution $B(\binom{x}{2}, p)$ when $x \geq 2$ and $|\mathcal{E}(V_1)| = 0$ when $x = 1$. $|\mathcal{E}(V_1, V_2)|$ follows a binomial distribution $B((n - x - 1)x, p)$. Under typical settings, $|\mathcal{E}(V_1)|$ is much smaller than $|\mathcal{E}(V_1, V_2)|$. For example, in $G(100, 0.05)$, given $|V_1| = 5$, we have $E[|\mathcal{E}(V_1)|] = 0.25$, whereas $E[|\mathcal{E}(V_1, V_2)|] = 94 \times 0.25$. Therefore, $2|\mathcal{E}(V_1)| + |\mathcal{E}(V_1, V_2)| \approx |\mathcal{E}(V_1)| + |\mathcal{E}(V_1, V_2)|$, which follows $B(\frac{1}{2}x(x-1) + (n-x-1)x, p)$. This binomial distribution is approximately normal since $(n - x - 1)x \cdot p$ and $(n - x - 1)x \cdot (1 - p)$ are both greater than 5 under typical settings in networks with practical-sizes (We usually have $np \geq \ln n$ in a connected graph $G(n, p)$ [5]). Therefore, $k(V_1)$ is approximately normal given $|V_1| = x$. Let μ and σ^2 denote the conditional mean and variance of $k(V_1)$ given $|V_1| = x$, respectively. They can be obtained by

$$\begin{aligned}
\mu &= x + x(x-1)p + (n-x-1)xp, \\
\sigma^2 &= (2x(x-1) + (n-x-1)x)p(1-p).
\end{aligned}$$

Such that $\Pr\{k(V_1) < \frac{k(V_0)}{\alpha-1} \mid |V_1| = x\}$ *in* (5.17) *can be approximated as*

$$\Pr\{k(V_1) < \frac{k(V_0)}{\alpha - 1} \mid |V_1| = x\} \approx \Phi(\frac{\frac{x}{\alpha-1} - \mu}{\sigma}). \tag{5.20}$$

After substituting (5.18) and (5.20) into (5.17), we obtain

$$E[|F_1|] \approx \sum_{x=1}^{n-1} x \cdot \binom{n-1}{x} p^x (1-p)^{n-1-x} \Phi(\frac{\frac{x}{\alpha-1} - \mu}{\sigma}).$$

By definition, $E[f_1] = \frac{1}{n}(1 + E[|F_1|])$*, which yields Theorem 1.*

Theorem 3.3.1 Consider a random single-node attack applied to $G(n,p)$. We assume,

1. We only consider the failures propagating in the forward direction; i.e., at step t, only the nodes in $V \setminus \cup_{d=0}^{t-1} V_d$ are considered as potential nodes to fail.

2. The set F_{t-1} is considered as a large virtual node that redistributes its load to its alive neighbors at step t with the rule defined in (3).

3. n is large enough such that the variance of $|\hat{V}_t|$ is small and $|\hat{V}_t|$ can be approximated by $E[|\hat{V}_t|]$.

4. $E[|\hat{V}_t| \mid |F_{t-1}| = E[|F_{t-1}|]]$ is applied to approximate $E[|\hat{V}_t|]$.

5. Given $|\hat{V}_t| = E[|\hat{V}_t|]$, $L_t(F_{t-1})$ and $(\alpha-1)L_0(\hat{V}_t)$ are independent and approximately normal. $L_t(F_{t-1})$ has conditional mean $\tilde{\mu} = E[|T_{t-1}|](n-1)p$ and unknown conditional variance $\tilde{\sigma}^2$. $(\alpha-1)L_0(\hat{V}_t)$ has conditional mean $\hat{\mu} = (\alpha-1)(n-1)E[|\hat{V}_t|]p$ and conditional variance $\hat{\sigma}^2 = (\alpha-1)^2(n-1)E[|\hat{V}_t|]p(1-p)$. $\Phi(\frac{\tilde{\mu}-\hat{\mu}}{\hat{\sigma}})$ is applied as an approximation of $\Pr\{L_t(F_{t-1}) > (\alpha-1)L_0(\hat{V}_t) \mid |\hat{V}_t| = E[|\hat{V}_t|]\}$.

Then an estimate of the average failure ratio $E[f_t]$ for step $t \geq 2$ is obtained recursively as

$$E[f_t] \approx \frac{1}{n}\Phi(\frac{\tilde{\mu}-\hat{\mu}}{\hat{\sigma}})E[|\hat{V}_t|] + E[f_{t-1}],$$

where

$$\begin{aligned} E[|\hat{V}_t|] &= (n - \sum_{d=0}^{t-1} E[|V_d|]) \cdot (1-(1-p)^{E[|F_{t-1}|]}), \\ E[|T_{t-1}|] &= nE[f_{t-1}], \\ E[|F_{t-1}|] &= n(E[f_{t-1}] - E[f_{t-2}]), \end{aligned}$$

$E[|V_0|] = 1$ by definition and $E[|V_d|],\ d \geq 1$ are given by Lemma 2.3.1.

proof 6 (Proof of Theorem 2) After step 1*,* $E[f_t]$ *depends on random variables* $|V_1|, |V_2|, \cdots, |V_{t-1}|$*, as well as* $k(V_1), k(V_2), \cdots, k(V_{t-1})$*. However, finding the joint distribution of all these random variables is very difficult. Therefore, we need to make several necessary simplifying assumptions and approximations to obtain a closed-form result, as listed in the theorem. In this proof, we will first use these assumptions and approximations to derive the approximation of* $E[f_t]$*, and then justify all the assumptions and approximations point-by-point in the end of the proof.*

According to the assumption 1): we only consider the failures propagating in the forward direction, i.e., at step t *only the nodes in* $V \setminus \cup_{d=0}^{t-1} V_d$ *are considered as potential nodes to fail, we define the set of target nodes* $\hat{V}_t$ *at step* t *as the nodes in* $V \setminus \cup_{d=0}^{t-1} V_d$ *connected to* F_{t-1}*, as illustrated in Fig. 5.2. Since each node in* $V \setminus \cup_{d=0}^{t-1} V_d$ *has probability* p *to be connected with a node in* F_{t-1} *independently,* $|\hat{V}_t|$ *obeys a binomial distribution* $B(n - \sum_{d=0}^{t-1} |V_d|, 1-(1-p)^{|F_{t-1}|})$*. The target nodes will receive redistributed load from* F_{t-1} *at step* t*.*

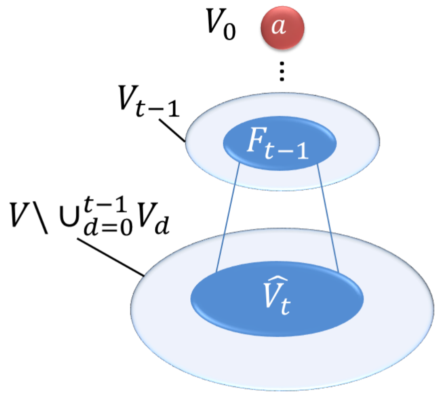

Figure 5.2: At step t, F_{t-1} redistributes its load to set $\hat{V}_t$. V_{t-1} is the set of nodes with a shortest distance $t-1$ to node a; F_{t-1} is the set of nodes failing at step $t-1$; $\hat{V}_t$ is the set of target nodes at step t, defined as the nodes in $V \setminus \cup_{d=0}^{t-1} V_d$ connected to F_{t-1}.

Furthermore, instead of analyzing the load redistribution for every node in F_{t-1}, we make the assumption 2): the set F_{t-1} is considered as a large virtual node that redistributes its load to its alive neighbors at step t. This assumption makes the analysis mathematically tractable. The problem now becomes "a single node redistributing its load to its neighbors". Similar to step 1*, by applying Corollary 1 to this equivalent setting, the failure condition for $\hat{V}_t$ can be found as*

$$L_t(F_{t-1}) + L_0(\hat{V}_t) \ > \ \alpha L_0(\hat{V}_t), \tag{5.21}$$

and $E[|F_t|]$ can be obtained as

$$E[|F_t|] \ = \ \sum_z z \Pr\left\{\hat{V}_t \, \textit{fails} \mid |\hat{V}_t| = z\right\} \Pr\{|\hat{V}_t| = z\}. \tag{5.22}$$

Recall our goal is to obtain $E[f_t]$ through

$$E[f_t] \ = \ E[|F_t|]/n + E[f_{t-1}],$$

where $E[f_{t-1}]$ is estimated in the previous step analysis. Now we aim to find $E[|F_t|]$ in (5.22)*. However, it is difficult to find the exact distribution of $|\hat{V}_t|$ in (5.22) as it requires the joint distribution of $|V_1|, |V_2|, \cdots, |V_{t-1}|$ and $k(V_1), k(V_2), \cdots, k(V_{t-1})$. According to the approximation 3): $E[|\hat{V}_t|]$ is applied to approximate $|\hat{V}_t|$, we have*

$$E[|F_t|] \;\approx\; \Pr\left\{\hat{V}_t \, \textit{fails} \mid |\hat{V}_t| = E[|\hat{V}_t|]\right\} E[|\hat{V}_t|], \tag{5.23}$$

where $E[|\hat{V}_t|]$ depends on random variable $|F_{t-1}|$:

$$E[|\hat{V}_t|] \;=\; \sum_y E[|\hat{V}_t| \mid |F_{t-1}| = y] \cdot \Pr\{|F_{t-1}| = y\}. \tag{5.24}$$

And by $\hat{V}_t$'s definition, we have

$$E[|\hat{V}_t| \mid |F_{t-1}| = y] \;=\; (n - \sum_{d=0}^{t-1} E[|V_d|]) \cdot (1 - (1-p)^y). \tag{5.25}$$

Now we aim to find $E[|\hat{V}_t|]$. To avoid finding $\Pr\{|F_{t-1}| = y\}$ and the summation in (5.24), we make the approximation 4): $E[|\hat{V}_t| \mid |F_{t-1}| = E[|F_{t-1}|]]$ is applied to approximate $E[|\hat{V}_t|]$. This approximation leads to

$$E[|\hat{V}_t|] \;\approx\; (n - \sum_{d=0}^{t-1} E[|V_d|]) \cdot (1 - (1-p)^{E[|F_{t-1}|]}), \tag{5.26}$$

where

$$E[|F_{t-1}|] \;=\; n(E[f_{t-1}]),$$

and $E[f_{t-1}]$ is given by the previous step analysis; $E[|V_0|] = 1$ by definition, and $E[|V_d|], \forall d \geq 1$ are given by Lemma 1. We now estimate the probability $\Pr\left\{\hat{V}_t \, \textit{fails} \mid |\hat{V}_t| = E[|\hat{V}_t|]\right\}$ in (5.23). According to the failure condition (5.21), $\Pr\left\{\hat{V}_t \, \textit{fails} \mid |\hat{V}_t| = E[|\hat{V}_t|]\right\}$ can be *obtained by*

$$\begin{aligned}
&\Pr\left\{\hat{V}_t \textit{ fails} \mid |\hat{V}_t| = E[|\hat{V}_t|]\right\} \\
&= \Pr\{L_t(F_t) > (\alpha-1)L_0(\hat{V}_t) \mid |\hat{V}_t| = E[|\hat{V}_t|]\}. \qquad (5.27)
\end{aligned}$$

According to assumption 5), the above probability can be approximated by

$$\Pr\{L_t(F_t) > (\alpha-1)L_0(\hat{V}_t) \mid |\hat{V}_t| = E[|\hat{V}_t|]\} \approx \Phi(\frac{\tilde{\mu}-\hat{\mu}}{\hat{\sigma}}), \qquad (5.28)$$

where

$$\begin{aligned}
\tilde{\mu} &= E[|T_{t-1}|](n-1)p, \\
\hat{\mu} &= (\alpha-1)(n-1)E[|\hat{V}_t|]p, \\
\hat{\sigma}^2 &= (\alpha-1)^2(n-1)E[|\hat{V}_t|]p(1-p).
\end{aligned}$$

By substituting (5.28) *and* (5.26) *into* (5.23)*, we find* $E[|F_t|]$*. Then* $E[f_t]$ *can be obtained as* $E[f_t] = E[|F_t|]/n + E[f_{t-1}]$*, which yields Theorem 2.*

Now we show the point-by-point justifications of all assumptions and approximations used:

1. *We only considered failures propagating in the forward direction, i.e., at step* t*, we only considered the nodes in* $V \setminus \cup_{d=0}^{t-1} V_d$ *as potential nodes to fail.*

   This assumption will be discussed in Section 3.4.

2. *The set* F_{t-1} *is considered as a large virtual node that redistributes its load to its alive neighbors at step* t *with the rule defined in (3).*

   First, this assumption is necessary to make the load redistribution after step 1 *math-*

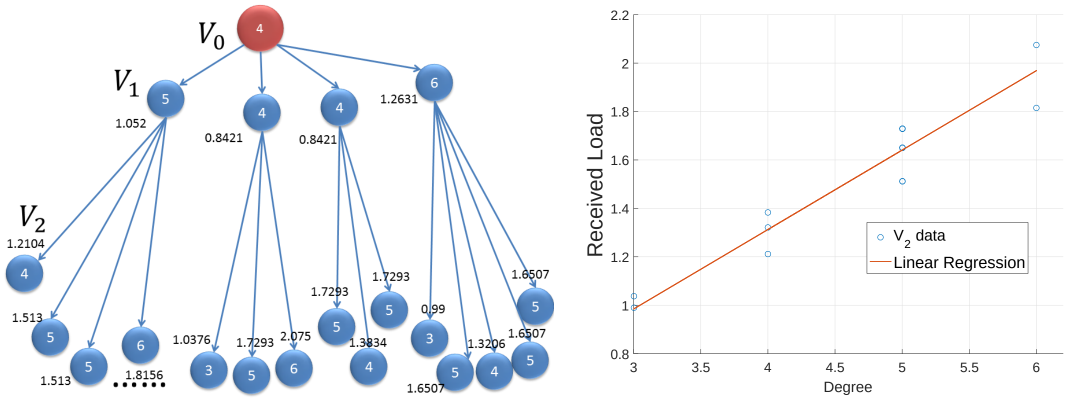

Figure 5.3: Example of load redistribution. White numbers located inside of circles are degrees, and numbers outside of circles are received load amounts.

ematically analyzable. Without this assumption, we would need to have the joint distribution of all node degrees in F_{t-1}*, as well as the link connections between* F_{t-1} *and* $\hat{V}_t$*, which is analytically complicated, especially when* t *is large.*

This assumption is appropriate in ER random graph because such graph is homogeneous by construction. In a typical realization of ER graph, the loads of nodes in F_{t-1} *have small variation. After the nodes in* F_{t-1} *distribute their loads to their alive neighbors according to their degrees, a neighbor node with a higher degree tends to receive more load, and vice versa. Such an example is shown in Fig. 5.3. The given partial network is a typical realization of ER random graph. Nodes in* V_1 *redistribute their loads to their neighbors in* V_2*. The scatter plot of the received load amounts and degrees in* V_2 *is also shown in Fig. 5.3. The linear correlation of the received load amounts and degrees in* V_2 *is* 0.9556*, indicating a strong linear relationship, which matches the assumed case.*

3. n *is large enough such that the variance of* $|\hat{V}_t|$ *is small and* $|\hat{V}_t|$ *can be approximated by* $E[|\hat{V}_t|]$*.*

By definition of target nodes, we have

$$|\hat{V}_t| \sim B(n - \sum_{d=0}^{t-1} |V_d|, 1 - (1-p)^{|F_{t-1}|}),$$

with variance

$$\begin{aligned} Var &= (n - \sum_{d=0}^{t-1} |V_d|)(1 - (1-p)^{|F_{t-1}|})(1-p)^{|F_{t-1}|} \\ &= E[|\hat{V}_t|] \cdot (1-p)^{|F_{t-1}|}. \end{aligned}$$

For $t \geq 2$ *and before the stage of steady state, we have* $|F_{t-1}| \gg 0$, $(1-p)^{|F_{t-1}|} \approx 0$ *and* $Var \approx 0$ *under typical settings of finite networks. For example, given* $p = 0.06$, $|F_{t-1}| = 100$ *and* $E[|\hat{V}_t|] = 40$, $Var = 0.0822$*. With Chebyshev's inequality, we have* $\Pr\{||\hat{V}_t| - 40| > 3 \times \sqrt{0.0822}\} \leq \frac{1}{9}$*. We can see that* $|\hat{V}_t|$ *stays close to its mean with high probability such that it can be approximated by its mean, as long as network' size is large enough (still finite). In addition, for the asymptotic case, we also have* $(1-p)^{|F_{t-1}|} \to 0$ *and* $Var \to 0$ *with* $|F_{t-1}| \to \infty$.

4. $E[|\hat{V}_t| \mid |F_{t-1}| = E[|F_{t-1}|]]$ *is applied to approximate* $E[|\hat{V}_t|]$.

   For notational simplicity, let $Y = |F_{t-1}|$*. Now our goal is to show that* $E[|\hat{V}_t|] \approx E[|\hat{V}_t| \mid Y = E[Y]]$*. According to* (5.24)*, we have*

$$E[|\hat{V}_t|] = \sum_y E[|\hat{V}_t| \mid Y = y] P(Y = y), \tag{5.29}$$

   where $E[|\hat{V}_t| \mid Y = y]$ *can be rewritten as a function of* y*:*

$$E[|\hat{V}_t| \mid Y = y] = c(1 - (1-p)^y) := f(y), \tag{5.30}$$

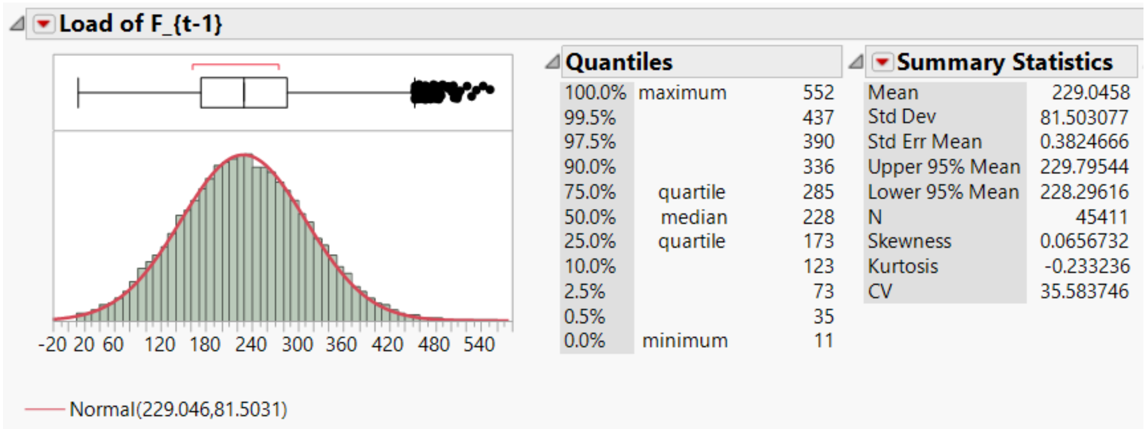

Figure 5.4: Histogram of $L_t(F_{t-1})$ and fit normal distribution. Skewness and kurtosis indicate that this distribution is quite symmetric, not heavily tailed.

with $c = (n - \sum_{d=0}^{t-1} E[|V_d|])$ *as a constant. Combining* (5.29) *with* (5.30), *we have*

$$E[|\hat{V}_t|] \;=\; E[f(Y)]. \tag{5.31}$$

We now aim to show that $E[f(Y)] \approx f(E[Y])$. *First we look at the derivative of* $f(y)$*:*

$$f'(y) \;=\; -c(1-p)^y \ln(1-p),$$

where c *is a positive constant,* $(1-p)^y < 1$*, and* $\ln(1-p) \approx 0$ *under typical settings in finite networks (where* p *is a small number). So* $f'(y)$ *is a very small positive number. For example, given* $c = 10, p = 0.08$*, and* $y = 30$*, we have* $f'(y) = -10 \cdot 0.92^{30} \cdot \ln(0.92) = 0.0683$*, which is close to zero. Since* $f'(y)$ *is close to zero and* $f(y)$ *is approximately constant over* y*,* $E[f(Y)]$ *can be approximated by:*

$$E[f(Y)] \;\approx\; f(y^*),$$

where y^* is an arbitrary point within $f(y)$'s domain. Let us pick $y^* = E[Y]$ such that we have $E[f(Y)] \approx f(E[Y])$, where $E[Y]$ is obtained from the previous step analysis. According to the simulation result, Y usually has a symmetric distribution and $E[Y]$ is the median of Y. Since $f(Y)$ is a monotonic increasing function, $f(E[Y])$ must be the median of $f(Y)$. Thus $f(E[Y])$ is a reasonable approximation of $E[f(Y)]$. Based on (5.31)*, we have*

$$E[|\hat{V}_t|] \ = \ E[f(Y)] \ \approx \ f(E[Y]).$$

According to the definition of $f(y)$ in (5.30)*, we have*

$$\begin{aligned} E[|\hat{V}_t|] \ \approx \ f(E[Y]) \ = \ E[X \mid Y = E[Y]] \\ = E[|\hat{V}_t| \mid |F_{t-1}| = E[|F_{t-1}|]]. \end{aligned}$$

5. *Given $|\hat{V}_t| = E[|\hat{V}_t|]$, $L_t(F_{t-1})$ and $(\alpha - 1)L_0(\hat{V}_t)$ are independent and approximately normal. $L_t(F_{t-1})$ has conditional mean $\tilde{\mu} = E[|T_{t-1}|](n-1)p$ and unknown variance. $(\alpha - 1)L_0(\hat{V}_t)$ has conditional mean $\hat{\mu} = (\alpha - 1)(n - 1)E[|\hat{V}_t|]p$ and conditional variance $\hat{\sigma}^2 = (\alpha-1)^2(n-1)E[|\hat{V}_t|]p(1-p)$. $\Phi(\frac{\tilde{\mu}-\hat{\mu}}{\hat{\sigma}})$ is applied as an approximation of $\Pr\{L_t(F_{t-1}) > (\alpha - 1)L_0(\hat{V}_t) \mid |\hat{V}_t| = E[|\hat{V}_t|]\}$.*

   For notational simplicity, let R_1 and R_2 denote random variables $L_t(F_{t-1})$ and $(\alpha-1)L_0(\hat{V}_t)$, conditional on $|\hat{V}_t| = E[|\hat{V}_t|]\}$, respectively. R_1 has mean $\tilde{\mu}$ and variance $\tilde{\sigma}^2$. R_2 has mean $\hat{\mu}$ and variance $\hat{\sigma}^2$. Thus, obtaining the probability $\Pr\{L_t(F_{t-1}) > (\alpha - 1)L_0(\hat{V}_t) \mid |\hat{V}_t| = E[|\hat{V}_t|]\}$ is equivalent to obtaining $\Pr\{R_1 > R_2\}$.

   First we look at R_1, which does not depend on $|\hat{V}_t|$ and it equals the summation over the initial loads in T_{t-1}. Since all the nodes' initial load amounts are i.i.d. random variables with distribution $B(n-1,p)$, we have $R_1 \sim B((n-1)|T_{t-1}|, p)$.

However, $|T_{t-1}|$ is itself a random variable with an unknown distribution. Only $E[|T_{t-1}|]$ is obtained from the previous step's analysis. Thus, $\tilde{\mu} = E[|T_{t-1}|](n-1)p$ is known, while $\tilde{\sigma}^2$ and the distribution of R_1 remain unknown. To get an idea of what distribution R_1 looks like, we show a histogram of $L_t(F_{t-1})$ from the simulation results in Fig. 5.4, which has the following settings: $n = 100, p = 0.06, \alpha = 1.1, t = 4$, and sample size is $50,000$. The case where no failures are triggered at step 1 was excluded in the histogram since no failures will happen at step 2 either in this scenario. From Fig. 5.4, we see that R_1 is approximately normal.

Then we look at R_2. Given that $|\hat{V}_t| = E[|\hat{V}_t|]$, we have $R_2 \sim B((n-1)E[|\hat{V}_t|], p)$, which is approximately normal because $E[|\hat{V}_t|] \gg 0$ before the steady state, such that $(n-1)E[|\hat{V}_t|]p$ and $(n-1)E[|\hat{V}_t|](1-p)$ are both greater than 5 under typical settings in a practical-size network. Given $\alpha - 1$ is a constant, $(\alpha - 1)L_0(\hat{V}_0)$ is also approximately normal. R_2's mean $\hat{\mu}$ and variance $\hat{\sigma}^2$ can be obtained as $\hat{\mu} = (\alpha-1)(n-1)E[|\hat{V}_t|]p$ and $\hat{\sigma}^2 = (\alpha-1)^2(n-1)E[|\hat{V}_t|]p(1-p)$.

Note R_1 and R_2 are independent because $L_0(\hat{V}_t)$ and $L_t(F_{t-1})$ will not affect each other's distribution with given $|\hat{V}_t|$. Then,

$$\begin{aligned} \Pr\{R_1 > R_2\} &= \Pr\{R_2 - R_1 < 0\} \\ &\approx \Phi(\frac{\tilde{\mu} - \hat{\mu}}{\sqrt{\tilde{\sigma}^2 + \hat{\sigma}^2}}), \end{aligned} \tag{5.32}$$

where $\tilde{\sigma}^2$ remains unknown. We notice that $\Pr\{R_1 > R_2\}$ does not depend on $\tilde{\sigma}^2$ in the following three cases:

(a) $\tilde{\mu} - \hat{\mu} \gg 0$ *and* (5.32) ≈ 1*,*

(b) $\tilde{\mu} - \hat{\mu} \ll 0$ *and* (5.32) ≈ 0*,*

(c) $\tilde{\mu} - \hat{\mu} = 0$ *and* (5.32) $= 0.5$*.*

Recall the physical meaning of $\Pr\{R_1 > R_2\} = \Pr\left\{\hat{V}_t \textit{ fails} \mid |\hat{V}_t| = E[|\hat{V}_t|]\right\}$ *in* (5.27). *The above cases a) and b) are dominant in a cascade since the threshold behavior (to be discussed in Section 3.4 and Section 4.) subjects to either "almost" all the nodes in* $\hat{V}_t$ *die or "almost" all of them survive. Therefore, we apply* $\Phi(\frac{\tilde{\mu}-\hat{\mu}}{\hat{\sigma}})$ *to approximate* $\Phi(\frac{\tilde{\mu}-\hat{\mu}}{\sqrt{\tilde{\sigma}^2+\hat{\sigma}^2}})$*, which has a good accuracy in the above three case. The accuracy of the above assumptions is further validated by Fig. 4.*

Theorem 4.1.1 Consider a random single-node attack applied to $G(n, p)$. Under the same assumption of Theorem 1. The critical value of α for the first step, i.e., $\alpha_{c,1}$, can be obtained by

$$\alpha_{c,1} \approx 1 + \frac{1}{1 + (n-2)p}.$$

proof 7 (Proof of Theorem 3) From Lemma 2, the nodes in V_1 *either all fail or all survive, we have*

$$\Pr\{v \textit{ fails} \mid v \in V_1\} = \Pr\{V_1 \textit{ fails}\}.$$

Then finding α *such that* $\Pr\{v \textit{ fails} \mid v \in V_1\} = \frac{1}{2}$*, i.e.,* $\alpha_{c,1}$*, is equivalent to finding* α *such that* $\Pr\{V_1 \textit{ fails}\} = \frac{1}{2}$.

According to Corollary 1 and (5.20) *in the proof of Theorem 1,* $\Pr\{V_1 \textit{ fails}\}$ *can be approximated as*

$$\Pr\{V_1 \textit{ fails}\} \approx \sum_{x=1}^{n-1} \binom{n-1}{x} p^x (1-p)^{n-1-x} \Phi\left(\frac{\frac{x}{\alpha-1} - \mu}{\sigma}\right), \tag{5.33}$$

where $\mu = x + x(x-1)p + (n-x-1)xp$*,* $\sigma^2 = (2x(x-1) + (n-x-1)x)p(1-p)$*,* $\binom{n-1}{x} = \frac{(n-1)!}{x!(n-x-1)!}$ *is the binomial coefficient,* $x = 1, \cdots, n-1$ *is the possible value of* $|V_1|$*, and* $\Phi(\cdot)$ *with is the CDF for the standard normal distribution. The approximation* (5.33) *was justified in the proof of Theorem 1. The summation* $\sum_{x=1}^{n-1}\binom{n-1}{x}p^x(1-p)^{n-1-x}$ *in* (5.33) *is* 1*, such that* $\Phi(\frac{x/(\alpha-1)-\mu}{\sigma}) = 0.5$ *is sufficient to guarantee that* $\Pr\{V_1 \textit{ fails}\} = 0.5$.

$$
\begin{aligned}
&\Phi(\frac{\frac{x}{\alpha-1}-\mu}{\sigma}) = 0.5,\\
\Leftrightarrow\ &\frac{x}{\alpha-1}-\mu = 0,\\
\Leftrightarrow\ &\frac{x}{\alpha-1}-(x+x(x-1)p+(n-x-1)xp) = 0,\\
\Leftrightarrow\ &\frac{1}{\alpha-1} = 1+(x-1)p+(n-x-1)p,\\
\Leftrightarrow\ &\frac{1}{\alpha-1} = 1+(n-2)p,\\
\Leftrightarrow\ &\alpha = \frac{1}{1+(n-2)p}+1.
\end{aligned}
\tag{5.34}
$$

We can see (5.34) *does not depend on* x*, i.e., under any possible values of* x, $\Phi(\frac{x/(\alpha-1)-\mu}{\sigma}) = 0.5$ *always holds when* $\alpha = 1+\frac{1}{1+(n-2)p}$*. Substitute it to* (5.33)*, we have* $\Pr\{V_1 \textit{ fails}\} \approx 0.5$*. By definition,* $\alpha_{c,1} \approx 1+\frac{1}{1+(n-2)p}$.

Theorem 4.1.2 Consider a random single-node attack applied to $G(n,p)$ with $\alpha \in (1, \frac{1}{(n-2)p}+1]$, and define a family of events: X_t="V_t fails at step t", $t \geq 1$. We assume that the set F_{t-1} can be considered as a large virtual node that redistributes its load to its alive neighbors at step t. Then the conditional probabilities of failure satisfy $\Pr\{X_t \mid X_{t-1}\} \geq \Pr\{X_1\}, \forall t > 1$.

proof 8 (Proof of Theorem 4) In ER random graph, the degree of each node has an independent distribution. To simplify the proof, we first define four types of distributions.

1. *Degree distribution of the initially attacked node* a*: Let* $E[k(a)] = \mu_1$ *and* $Var(k(a)) = \sigma_1^2$*. From the previous analysis, we know that* $k(a) \sim B((n-1),p)$*, which implies* $\mu_1 = (n-1)p$*, and* $\sigma_1^2 = (n-1)p(1-p)$.

2. *Degree distribution of an arbitrary node $e \in V_d$, $d > 0$: Let $E[k(e)] = \mu_2$, and $Var(k(e)) = \sigma_2^2$. We know e is connected to some nodes in the set V_{d-1}, such that its minimum degree is 1. We have $k(e) - 1 \sim B((n-2), p)$. This implies $\mu_2 = 1 + (n-2)p$ and $\sigma_2^2 = (n-2)p(1-p)$.*

3. *Degree distribution of V_1: Assuming $E[k(V_1)] = \mu_3$ and $Var[k(V_1)] = \sigma_3^2$, we have*

$$\begin{aligned}\mu_3 &= E_{|V_1|}[\sum_{v \in V_1} E[k(v)]] = E[|V_1|\mu_2] \\ &= (n-1)p[1+(n-2)p], \qquad (5.35)\\ \sigma_3^2 &= E_{|V_1|}[\sum_{v \in |V_1|} Var(k(v))] = E[|V_1|\sigma_2^2] \\ &= (n-1)p(n-2)p(1-p). \qquad (5.36)\end{aligned}$$

4. *Degree distribution of neighbor set $N(e)$, for arbitrary $e \in V_d,\ d > 0$: By letting $E[k(e)] = \mu_4$ and $Var[k(e)] = \sigma_4^2$, we have*

$$\begin{aligned}\mu_4 &= E_{|N(e)|}[\sum_{i \in N(e)} E[k(i)]] = E[|N(e)|\mu_2] \\ &= [1+(n-2)p]^2, \qquad (5.37)\\ \sigma_4^2 &= E_{|N(e)|}[\sum_{m \in |N(e)|} Var(k(m))] = E[|N(e)|\sigma_2^2] \\ &= [1+(n-2)p](n-2)p(1-p). \qquad (5.38)\end{aligned}$$

Now, we use these four distributions to prove the theorem. According to Corollary 1, the failure probability at the first step is

$$\Pr\{X_1\} = \Pr\{k(V_0) + (1-\alpha)k(V_1) > 0\}. \qquad (5.39)$$

For arbitrary $t > 0$, according to the assumption: the set F_{t-1} is considered as a large virtual node that redistributes its load to its alive neighbors at step t (justified in the proof of Theorem 2), given V_t fails at step t, the failure condition for V_{t+1} fails at step $t+1$ can be found as

$$L_t(V_t) + L_0(V_{t+1}) > \alpha L_0(V_{t+1}),$$

which is equivalent to

$$\sum_{d=0}^{t} k(V_d) + k(V_{t+1}) > \alpha k(V_{t+1}).$$

Such that the conditional probability $\Pr\{X_{t+1} \mid X_t\}$ can be obtained as

$$\begin{aligned}
&\Pr\{X_{t+1} \mid X_t\}, \\
=&\Pr\{\sum_{d=0}^{t} k(V_d) + (1-\alpha)k(V_{t+1}) \\
&> 0 \mid \sum_{d=0}^{t-1} k(V_d) > (\alpha-1)k(V_t)\}, \qquad (5.40) \\
&\xrightarrow{\sum_{d=0}^{t-1} k(V_d) > (\alpha-1)k(V_t)} \\
>&\Pr\{(\alpha-1)k(V_t) + k(V_t) + (1-\alpha)k(V_{t+1}) > 0\}, \qquad (5.41) \\
=&\Pr\{\alpha k(V_t) + (1-\alpha)k(V_{t+1}) > 0\}, \qquad (5.42) \\
&\xrightarrow{k(V_{t+1}) \leq \sum_{m\in V_t} k(N(m));\alpha\geq 1} \\
\geq&\Pr\{\alpha \sum_{m\in V_t} k(m) + (1-\alpha) \sum_{m\in V_t} k(N(m)) > 0\}, \qquad (5.43) \\
=&\Pr\{\alpha \frac{\sum_{m\in V_t} k(m)}{|V_t|} + (1-\alpha)\frac{\sum_{m\in V_t} k(N(m))}{|V_t|} > 0\}, \qquad (5.44) \\
&\xrightarrow{\alpha>1} \\
>&\Pr\{\frac{\sum_{m\in V_t} k(m)}{|V_t|} + (1-\alpha)\frac{\sum_{m\in V_t} k(N(m))}{|V_t|} > 0\}. \qquad (5.45)
\end{aligned}$$

Let γ denote $k(V_0) + (1-\alpha)k(V_1)$ in (5.39). Also, let β_1 denote $\alpha\frac{\sum_{m\in V_t} k(m)}{|V_t|} + (1-\alpha)\frac{\sum_{m\in V_t} k(N(m))}{|V_t|}$ in (5.44) and β_2 denote $\frac{\sum_{m\in V_t} k(m)}{|V_t|} + (1-\alpha)\frac{\sum_{m\in V_t} k(N(m))}{|V_t|}$ in (5.45). Our goal is to obtain the α interval where (5.40) is equal to or greater than $\Pr\{\gamma > 0\}$. In order to find this interval accurately, we look at both cases: $\Pr(\beta_1 > 0) \geq \Pr(\gamma > 0)$ and $\Pr(\beta_2 > 0) \geq \Pr(\gamma > 0)$. Note that in the expressions of β_1 and β_2, for all $m \in V_t$, $k(m)$ and $k(N(m))$ maintain distributions 2) and 4) defined in the beginning of this proof, respectively. Assuming $|V_t| = b$, we have $\frac{\sum_{m\in V_t} k(m)}{|V_t|}$ as the sample mean of b i.i.d. random variables with the distribution 2) and $\frac{\sum_{m\in V_t} k(N(m))}{|V_t|}$ as the sample mean of b i.i.d. random variables with distribution 4). Hence, the means of γ, β_1 and β_2 satisfy

$$\begin{aligned}
E[\gamma] &= \mu_1 + (1-\alpha)\mu_3 \\
&= (n-1)p + (1-\alpha)(n-1)p[1+(n-2)p], \qquad (5.46)\\
E[\beta_1] &= \alpha\mu_2 + (1-\alpha)\mu_4 \\
&= \alpha[1+(n-2)p] + (1-\alpha)[1+(n-2)p]^2, \qquad (5.47)\\
E[\beta_2] &= \mu_2 + (1-\alpha)\mu_4 \\
&= [1+(n-2)p] + (1-\alpha)[1+(n-2)p]^2. \qquad (5.48)
\end{aligned}$$

The variances of γ, β_1 and β_2 satisfy

$$\begin{aligned}
Var[\gamma] &= \sigma_1^2 + (1-\alpha)^2\sigma_3^2 = (n-1)p(1-p)+ \\
&\quad (1-\alpha)^2(n-1)p(n-2)p(1-p), \qquad (5.49)\\
Var[\beta_1] &= \alpha^2\frac{\sigma_2^2}{b} + (1-\alpha)^2\frac{\sigma_4^2}{b} = \alpha^2\frac{(n-2)p(1-p)}{b}+ \\
&\quad (1-\alpha)^2\frac{[1+(n-2)p](n-2)p(1-p)}{b}, \qquad (5.50)
\end{aligned}$$

$$Var[\beta_2] = \frac{\sigma_2^2}{b} + (1-\alpha)^2\frac{\sigma_4^2}{b} = \frac{(n-2)p(1-p)}{b} + (1-\alpha)^2\frac{[1+(n-2)p](n-2)p(1-p)}{b}. \tag{5.51}$$

From these equations, we see that $E[\beta_1] > E[\beta_2] > E[\gamma]$ *and* $Var(\beta_2) < Var(\gamma)$ *for all* b*.*

Since γ, β_1, β_2 *are linear combinations of random degrees of multiple nodes, we apply CLT to approximate their distribution with normal distributions. In order to find the* α *interval satisfying* $\Pr\{X_{t+1} \mid X_t\} > \Pr\{\gamma > 0\}$*, we need to consider the following two situations:*

1. $E[\gamma] \leq 0$ *and* $E[\beta_1] \geq 0$ *when* $\Pr(\beta_1 > 0) \geq \Pr(\gamma > 0)$*,*

2. $E[\beta_2] > E[\gamma] \geq 0$ *and* $Var[\beta_2] < Var(\gamma)$ *when* $\Pr(\beta_2 > 0) \geq \Pr(\gamma > 0)$*.*

In both situations, according to (5.40)-(5.45), we have $\Pr\{X_{t+1} \mid X_t\} > \Pr\{\gamma > 0\}$*. Situation 1 leads to* $\alpha \in [\frac{1}{1+(n-2)p}+1, \frac{1}{(n-2)p}+1]$*, and Situation 2 leads to* $\alpha \in (1, \frac{1}{1+(n-2)p}+1]$*. Putting the two situations together, we have* $\Pr\{X_{t+1} \mid X_t\} > \Pr\{X_1\}$ *when* $\alpha \in (1, \frac{1}{(n-2)p} + 1]$*.*

Theorem 4.2.1 Consider a random single-node attack applied to $G(n, p)$. Under the same assumption of Theorem 1. We also assume $\alpha_{h,1}$ and $\alpha_{l,1}$ are estimated given $|V_1| = (n-1)p$. The threshold interval $[\alpha_{l,1}, \alpha_{h,1}]$ for the first step of the cascade can be obtained as follows:

$$\alpha_{h,1} \approx 1 + \frac{1}{-c\sqrt{(p+\frac{n-3}{n-1})(1-p)} + 1 + (n-2)p},$$
$$\alpha_{l,1} \approx 1 + \frac{1}{c\sqrt{(p+\frac{n-3}{n-1})(1-p)} + 1 + (n-2)p},$$

where $c = 2$ and $c = 3$ lead to the 2σ and 3σ threshold intervals, respectively. Within the 2σ and 3σ threshold intervals, $E[f_1]$ drops 95% and 99% of its maximum value, respectively.

proof 9 (Proof of Theorem 5) According to the empirical rule for a normal distribution, two and three standard deviations from the mean account for about 95% *and* 99% *of the probabilities, respectively. For readability, we rewrite* (5.33) *here:*

$$\Pr\{V_1 \textit{ fails}\} \approx \sum_{x=1}^{n-1} \binom{n-1}{x} p^x (1-p)^{n-1-x} \Phi(\frac{\frac{x}{\alpha-1} - \mu}{\sigma}), \tag{5.52}$$

where $\mu = x + x(x-1)p + (n-x-1)xp$*,* $\sigma^2 = (2x(x-1) + (n-x-1)x)p(1-p)$*,* $x = 1, \cdots, n-1$ *is the possible value of* $|V_1|$*.* $\Pr\{V_1 \textit{ fails}\}$ *is a weighted summation of the standard normal CDF function* $\Phi(\frac{x/(\alpha-1)-\mu}{\sigma})$ *over* $x = 1, \cdots, n-1$*. The normal approximation in* (5.52) *was justified in the proof of Theorem 1.*

Notice that the summation $\sum_{x=1}^{n-1} \binom{n-1}{x} p^x (1-p)^{n-1-x} = 1$ *in* (5.52)*. By definition of threshold interval,* $\alpha_{l,1}$ *is the* α *value such that* $(\frac{\frac{x}{\alpha-1}-\mu}{\sigma}) = c$*, and* $\alpha_{h,1}$ *is the* α *value such that* $(\frac{\frac{x}{\alpha-1}-\mu}{\sigma}) = -c$*, where* $c = 2$ *or* $c = 3$*, represents the* 2σ *or* 3σ *threshold interval, respectively. First, we find the higher boundary* $\alpha_{h,1}$*.*

$$\begin{aligned}
(\frac{\frac{x}{\alpha-1} - \mu}{\sigma}) &= -c, \\
\Leftrightarrow \frac{x}{\alpha-1} - \mu &= -c\sigma, \\
\Leftrightarrow \frac{x}{\alpha-1} &- (x + x(x-1)p + (n-x-1)xp) \\
&= -c\sqrt{(2x(x-1) + (n-x-1)x)p(1-p)}, \\
\Leftrightarrow \frac{1}{\alpha-1} &- (1 + (n-2)p) = -c\sqrt{(1 + nx^{-1} - 3x^{-1})p(1-p)},
\end{aligned} \tag{5.53}$$

For notational simplicity, let $\eta = 1 + (n-2)p$ *and* $\delta = \sqrt{(1 + nx^{-1} - 3x^{-1})p(1-p)}$*.*

Then, (5.53) *leads to*

$$\alpha_{h,1} \; = \; 1 + \frac{1}{-c\delta + \eta}. \tag{5.54}$$

Similarly, the lower boundary $\alpha_{l,1}$ can be obtained as

$$\alpha_{l,1} \; = \; 1 + \frac{1}{c\delta + \eta}. \tag{5.55}$$

In the above expressions, δ depends on x. Our goal is to eliminate x from (5.54) *and* (5.55)*, such that* (5.52) *would be a formula without the weighted sum over all x values. Notice that the binomial distribution in* (5.52) *has a bounded two-side tail probabilities $\sum_{x=1}^{2}\binom{n-1}{x}p^x(1-p)^{n-1-x}$ and $\sum_{x=3np}^{n-1}\binom{n-1}{x}p^x(1-p)^{n-1-x}$ [?], which are both negligible in a connected ER graph with practical-sizes. So we focus on the domain $x \in [3, 3np]$. First we show $\eta > 4\delta$ (to be used later).*

$$\begin{aligned}
&\eta^2 - 16\delta^2\\
\xrightarrow{1-p<1} &> 1 + (n-2)^2p^2 + 2(n-2)p - 16(1 + \frac{1}{x}(n-3))p\\
\xrightarrow{x\geq 3} &\geq 1 + (n-2)^2p^2 + 2(n-2)p - 16(1 + \frac{1}{3}(n-3))p\\
&= 1 + (n-2)^2p^2 + 2(n-2)p - 16(\frac{2}{3} + \frac{1}{3}(n-2))p\\
&= 1 + (n-2)^2p^2 - \frac{10}{3}(n-2)p - \frac{32}{3}p\\
&= ((n-2)p - \frac{5}{3})^2 - \frac{16}{9} - \frac{32}{3}p,
\end{aligned} \tag{5.56}$$

which is positive under typical settings ($np > \log n$, p is small) in a connected ER graph, such that $\eta > 4\delta$.

To see how $\alpha_{h,1}$ and $\alpha_{l,1}$ changes with x, we first consider the partial derivative of $\alpha_{h,1}$

over x. According to (5.54)*, we have*

$$
\begin{aligned}
&|\frac{\partial \alpha_{h,1}}{\partial x}| \\
=&\frac{1}{(\eta - c\delta)^2} \cdot c\sqrt{p(1-p)}\frac{1}{2}(1 + nx^{-1} - 3x^{-1})^{-\frac{1}{2}}(nx^{-2} - 3x^{-2}) \\
=&\frac{\delta}{(\eta - c\delta)^2} \cdot \frac{(n-3)x^{-1}}{1+(n-3)x^{-1}} \cdot \frac{1}{x} \cdot \frac{c}{2}.
\end{aligned}
\tag{5.57}
$$

Here we show that $|\frac{\partial \alpha_{h,1}}{\partial x}|$ is a very small under typical settings in networks with practical-sizes. Assuming $c = 3$, we have $\frac{(n-3)x^{-1}}{1+(n-3)x^{-1}} < 1$, $\frac{1}{x} \cdot \frac{c}{2} \leq \frac{1}{2}$ and $\frac{\delta}{(\eta - c\delta)^2} < \frac{\frac{1}{4}\eta}{(\eta - \frac{3}{4}\eta)^2} = \frac{4}{\eta} = \frac{4}{1+(n-2)p}$ (since $\delta < \frac{1}{4}\eta$), which is usually smaller than 1 under typical settings. Thus, $|\frac{\partial \alpha_{h,1}}{\partial x}|$ is usually smaller than 0.5. When $c = 2$, it can be shown that $|\frac{\partial \alpha_{h,1}}{\partial x}|$ becomes even smaller than the case when $c = 3$, using a similar method. For $\alpha_{l,1}$, we have

$$
|\frac{\partial \alpha_{l,1}}{\partial x}| = \frac{\delta}{(\eta + c\delta)^2} \cdot \frac{(n-3)x^{-1}}{1+(n-3)x^{-1}} \cdot \frac{1}{x} \cdot \frac{c}{2}, \tag{5.58}
$$

which is smaller than $|\frac{\partial \alpha_{h,1}}{\partial x}|$, since $(\eta + c\delta)^2 > (\eta - c\delta)^2$. For example, given $n = 100, p = 0.05, x = 5, c = 2$, we have $|\frac{\partial \alpha_{h,1}}{\partial x}| = 0.0121$ and $|\frac{\partial \alpha_{l,1}}{\partial x}| = 0.003$. Since $\frac{\partial \alpha_{h,1}}{\partial x}$ and $\frac{\partial \alpha_{l,1}}{\partial x}$ are always very small, i.e., $\alpha_{h,1}$ and $\alpha_{l,1}$ are approximately constant over $x \in [3, 3np]$, we can eliminate x by replacing it with its mean $E[|V_1|] = (n-1)p$ (given by the proof of Theorem 1). Substituting $x = (n-1)p$ into (5.54) *and* (5.55)*, it leads to the estimates of $\alpha_{l,1}$ and $\alpha_{h,1}$:*

$$
\alpha_{h,1} \;\approx\; 1 + \frac{1}{-c\sqrt{(p + \frac{n-3}{n-1})(1-p)} + 1 + (n-2)p}, \tag{5.59}
$$

$$
\alpha_{l,1} \;\approx\; 1 + \frac{1}{c\sqrt{(p + \frac{n-3}{n-1})(1-p)} + 1 + (n-2)p}, \tag{5.60}
$$

where c can be 2 or 3 based on the 2σ or 3σ criteria, respectively.